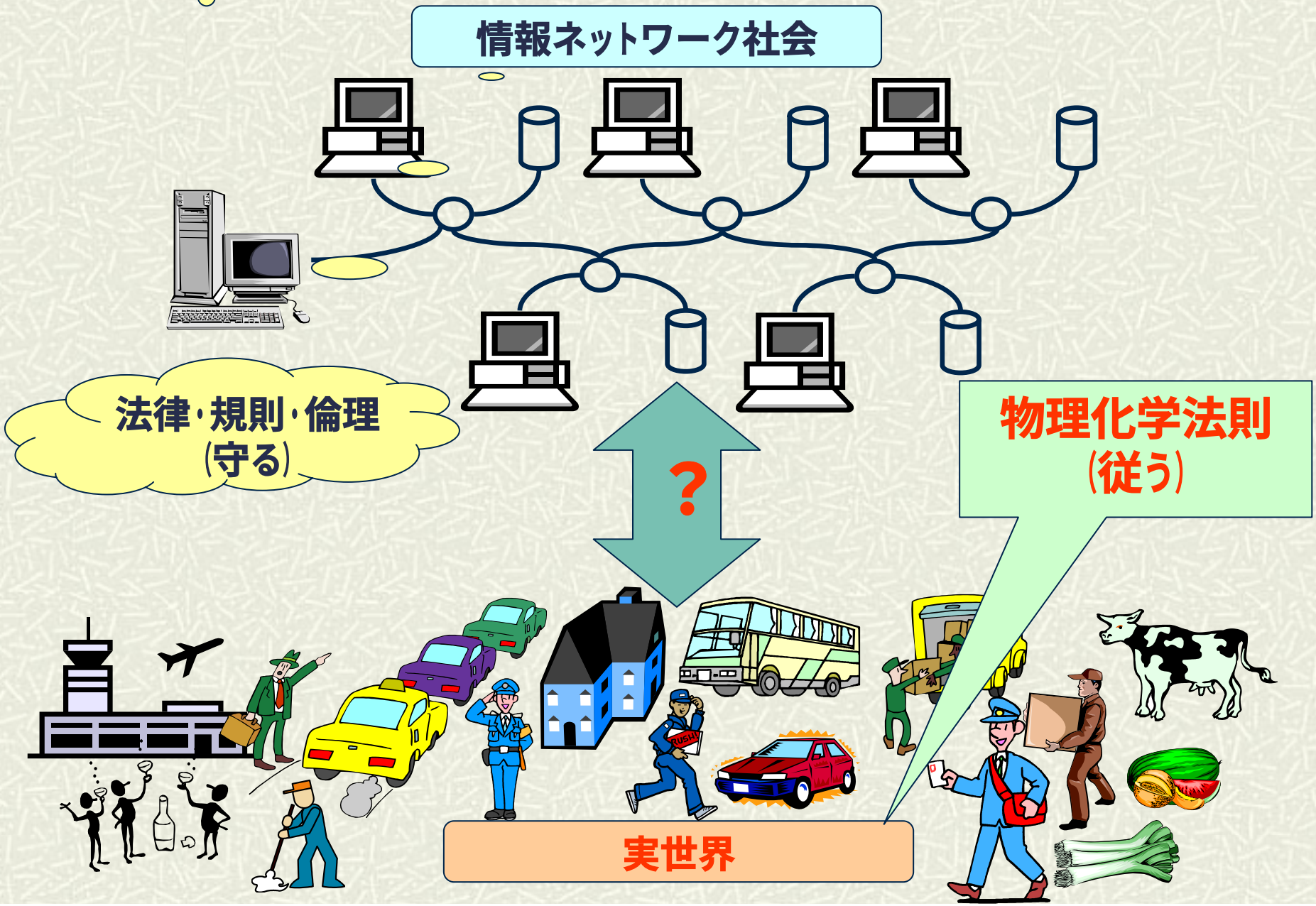
コンピュータ・ビジョン
実世界画像処理による
情報ネットワーク社会と実世界との統合

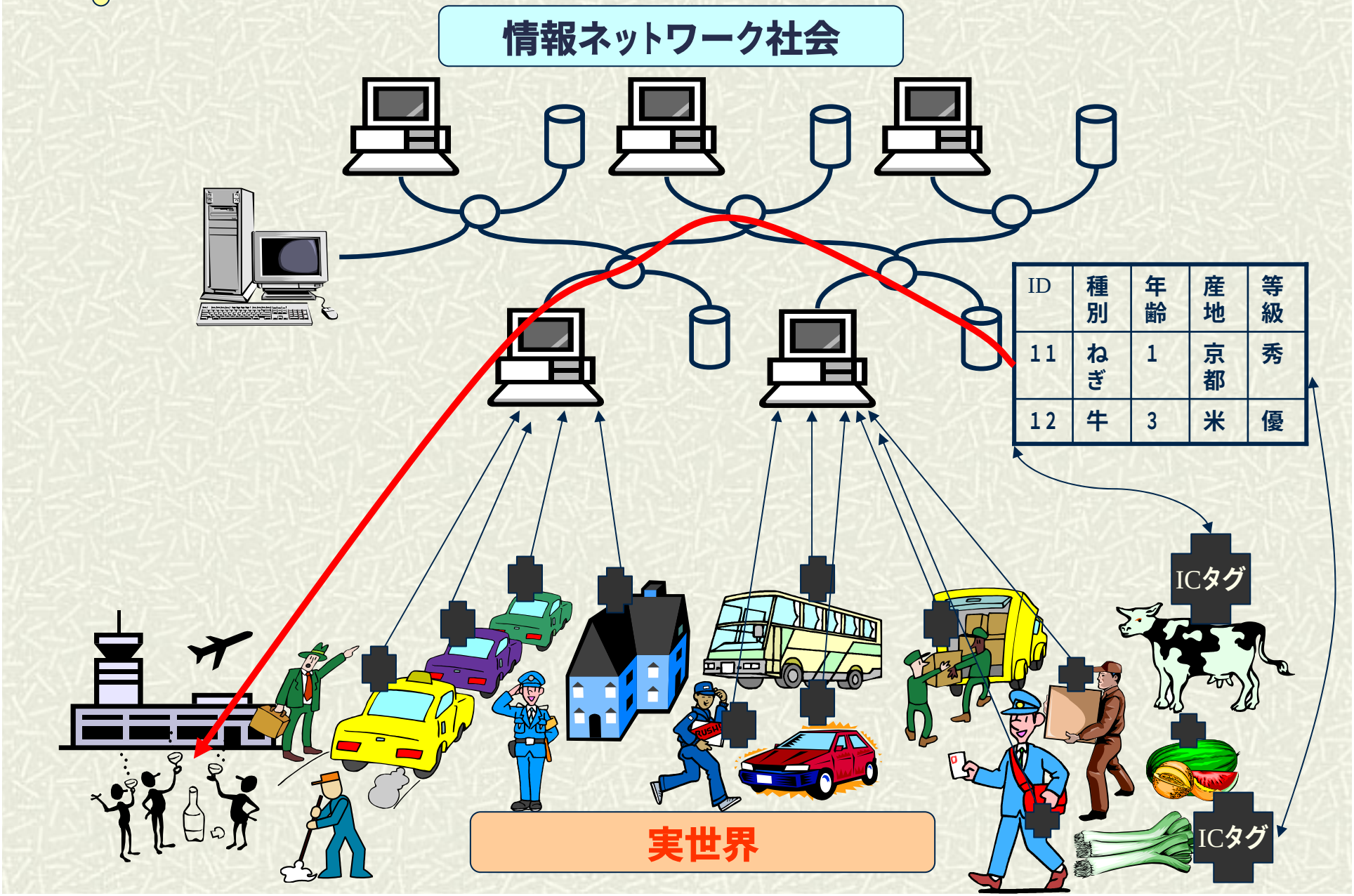
2008年1月24日

京都大学大学院　情報学研究科
松山　隆司

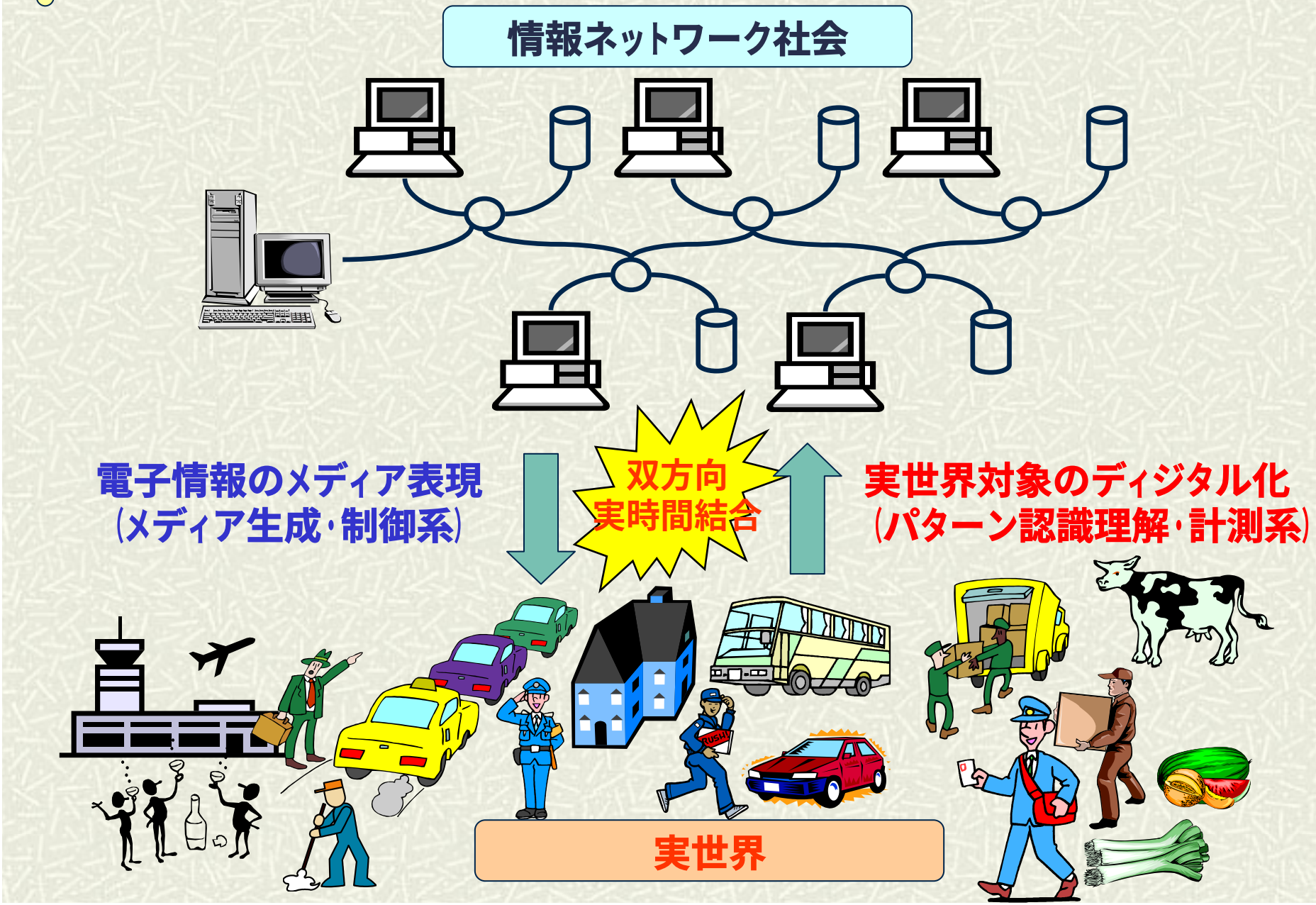
21世紀社会の構造



(I) ICタグによる両世界の対象の対応付け



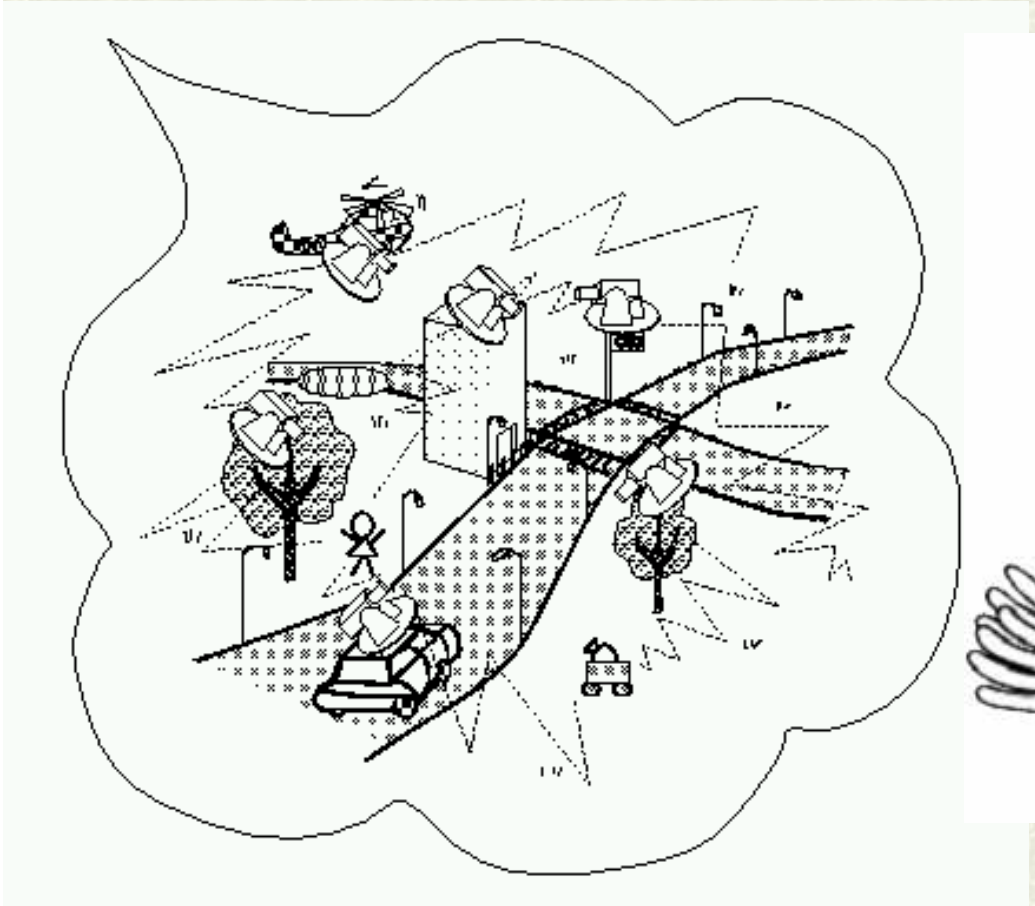
(II) マルチメディア情報処理による実時間双方向実時間結合

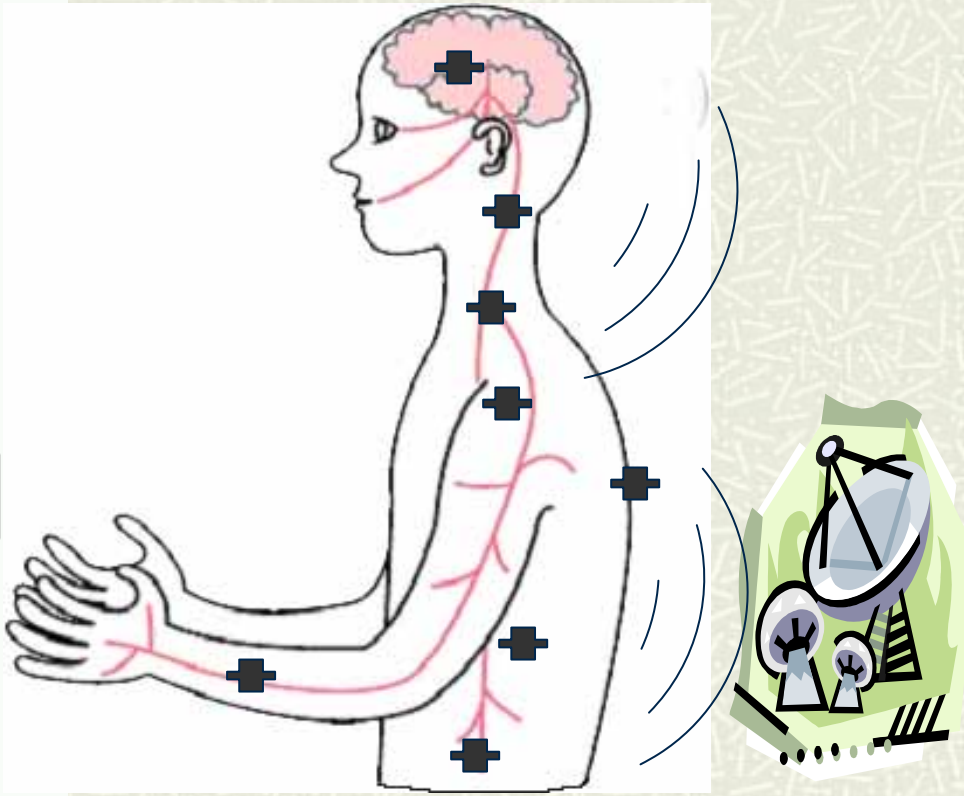


実世界埋め込み型センサネットワーク

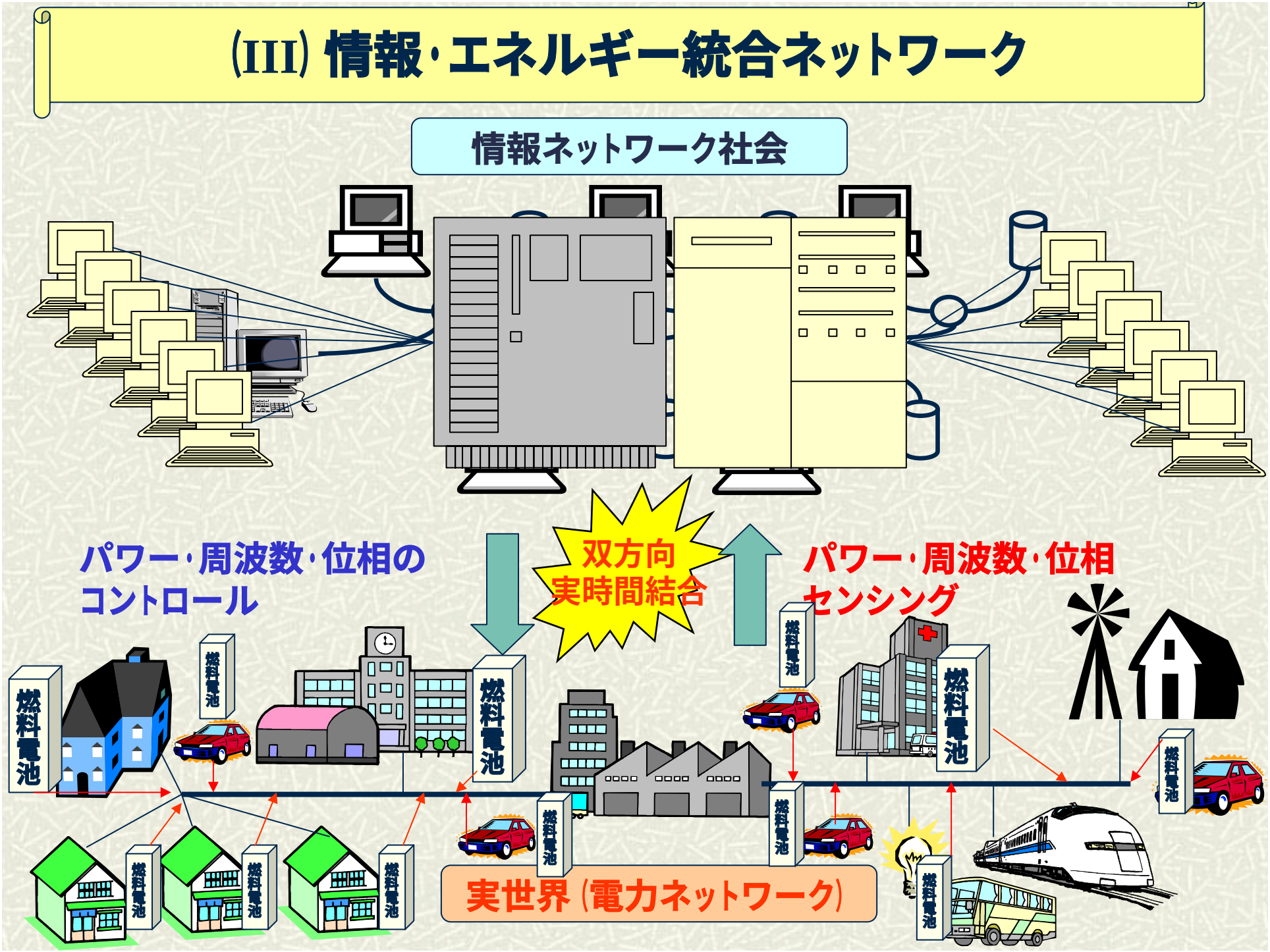
分散協調視覚システム
環境埋め込み型アクティブ・センサ・ネットワーク

体内埋め込み型ナノ・デバイス



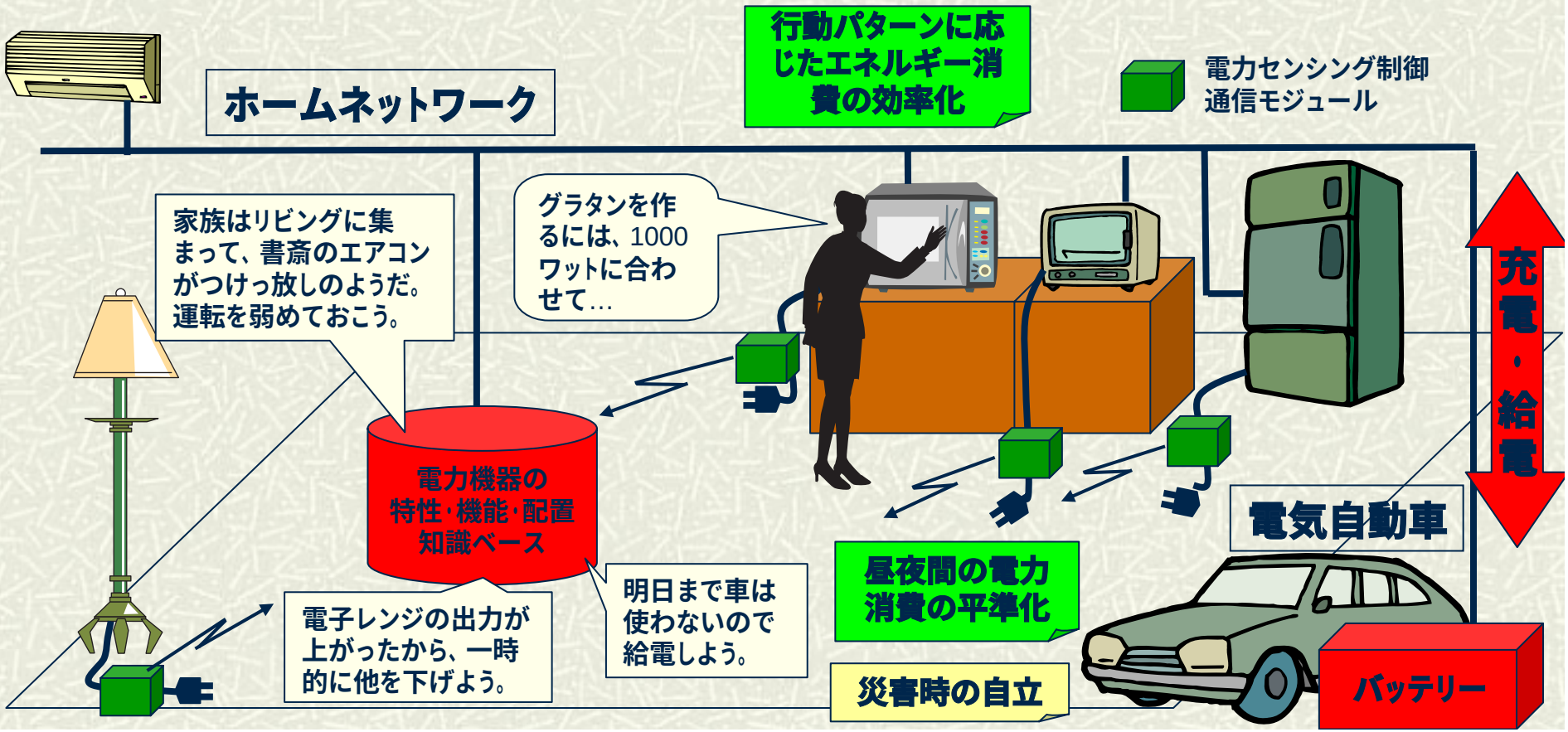


(III) 情報・エネルギー統合ネットワーク

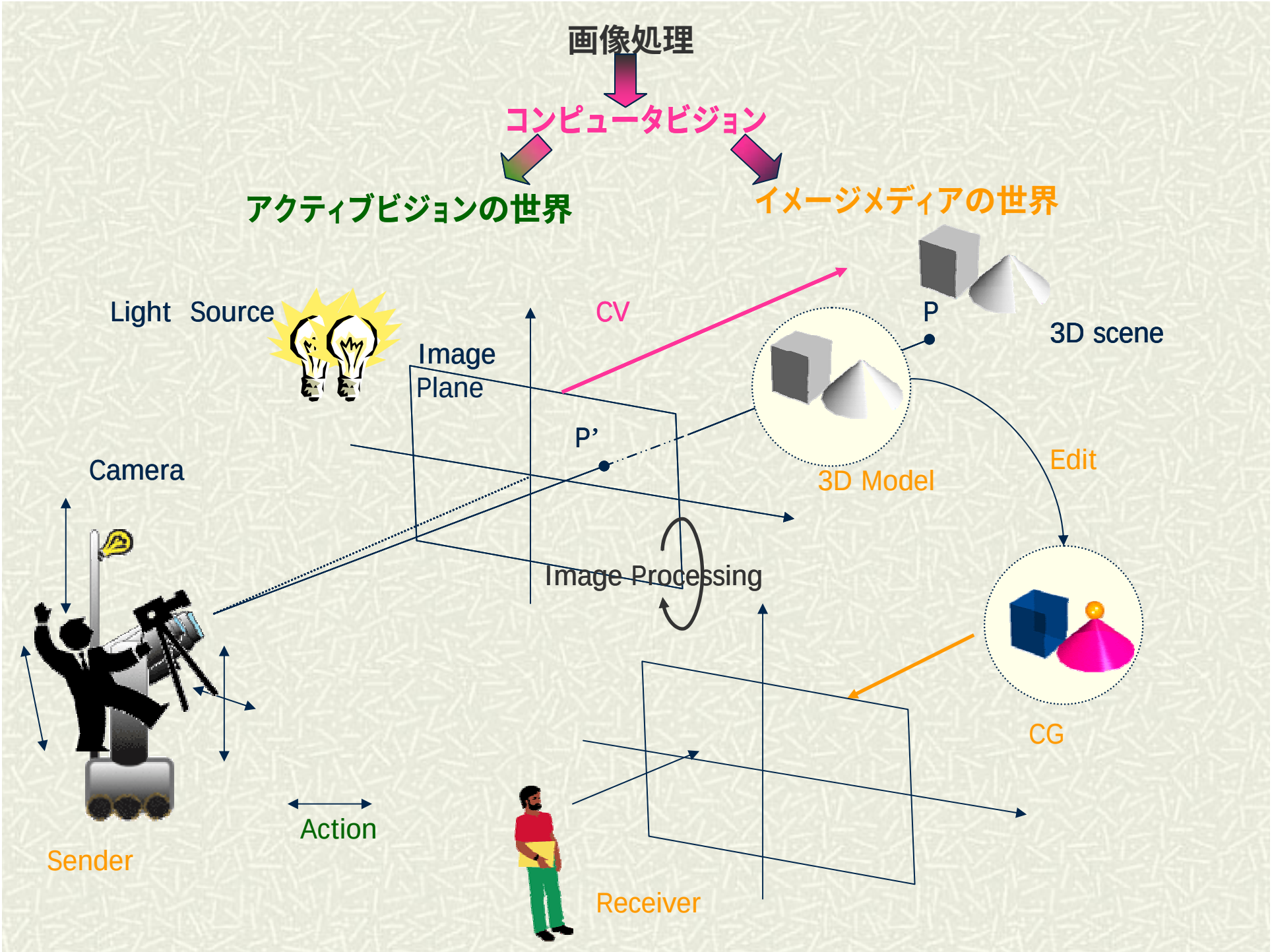


安全･安心なエコライフの実現
(電気自動車のエコライフ･エコ社会活用)

家庭内の電気機器に電力センシング･制御･通信モジュールを付加し、**電気自動車を活用した蓄･給電**も含め、各電気機器をネットワーク経由で制御する。人間の生活パターンの認識、漏電検知や**災害時の自立**生活支援による安心･安全な生活、**電力消費平準化**によるエコ生活支援を行う。



コンピュータビジョンの世界

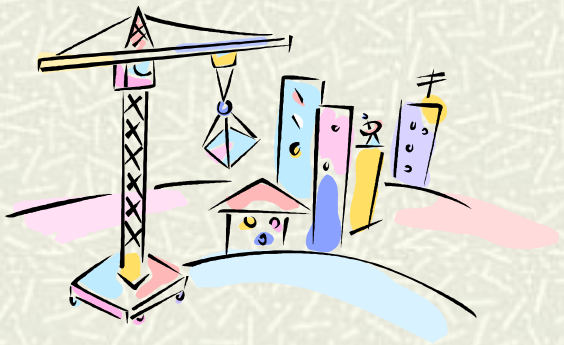

画像処理
コンピュータビジョン
アクティブビジョンの世界
イメージメディアの世界
Light Source
CV
P
3D scene
Image
Plane
P'
Camera
3D Model
Edit
Image Processing
CG
Action
Sender
Receiver

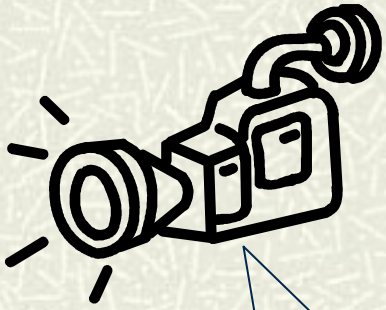
実世界画像処理
(計測装置としてのカメラ)

家庭用カメラ

「人が観るため」の画像

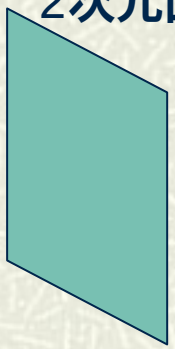
3次元シーン





輪郭強調
ガンマ補正
色調補正
・・・

2次元画像



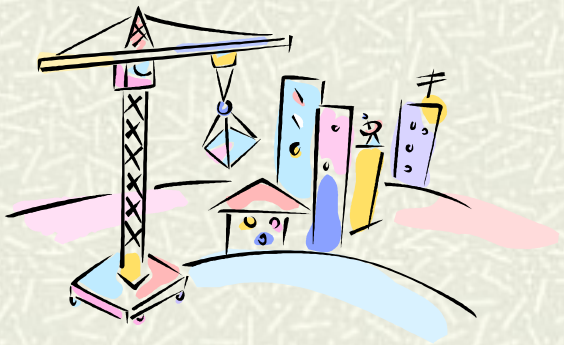
肌色がきれい
輪郭がくっきり
CRTで観やすい

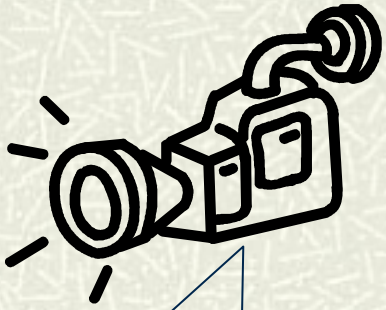
11

コンピュータビジョン用（産業用）

- 「素性の分かった」画像

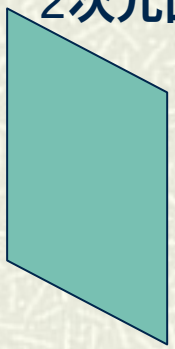
3次元シーン





輪郭強調 OFF
ガンマ補正 OFF
色調補正 OFF
・・・

2次元画像

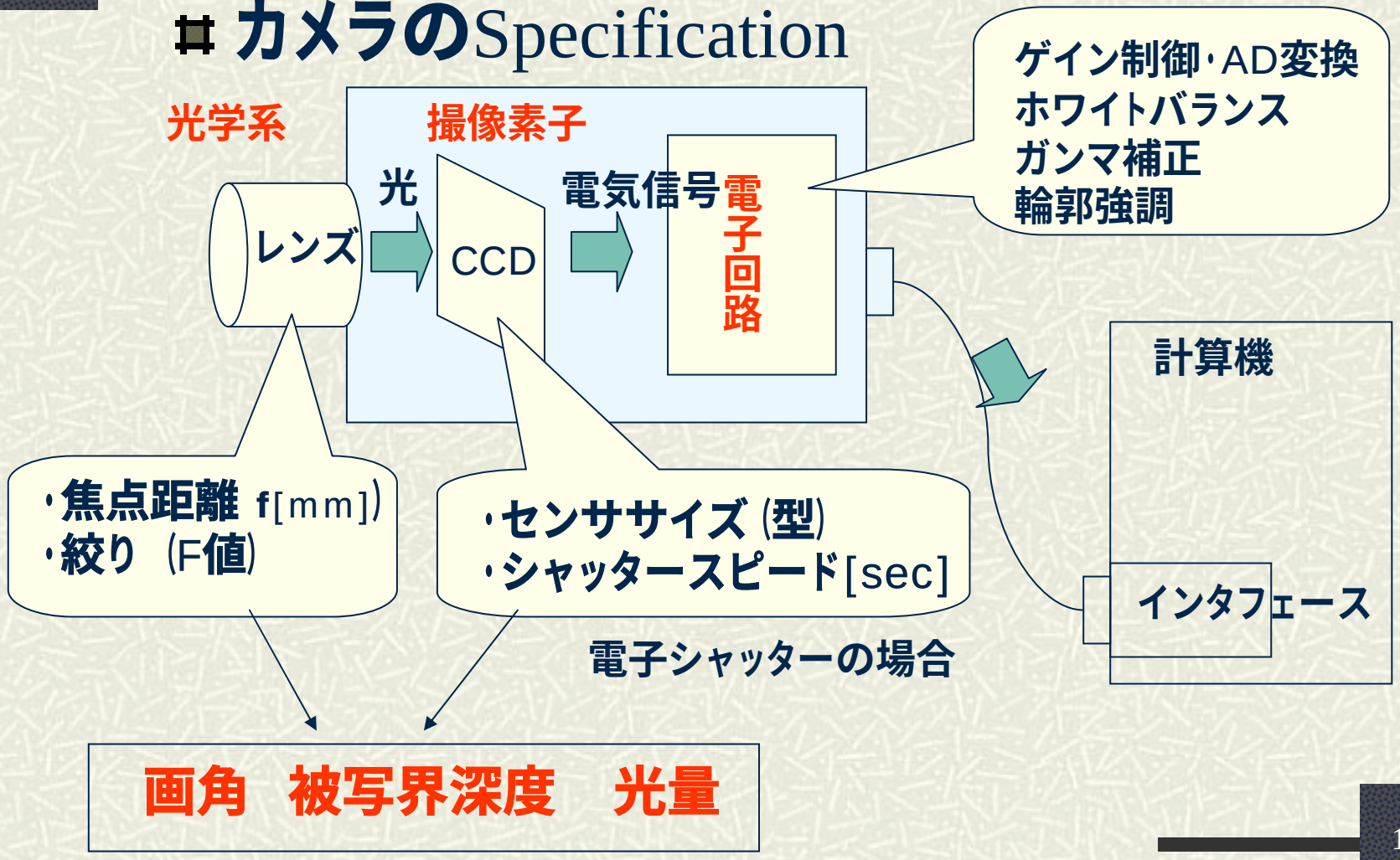


見た目には美しくなくても

12

画像入力プロセス

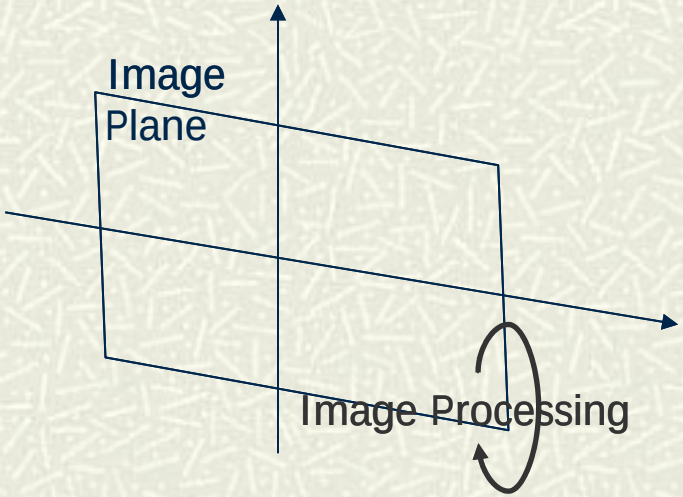
カメラのSpecification



13

画像処理と画像認識

画像処理



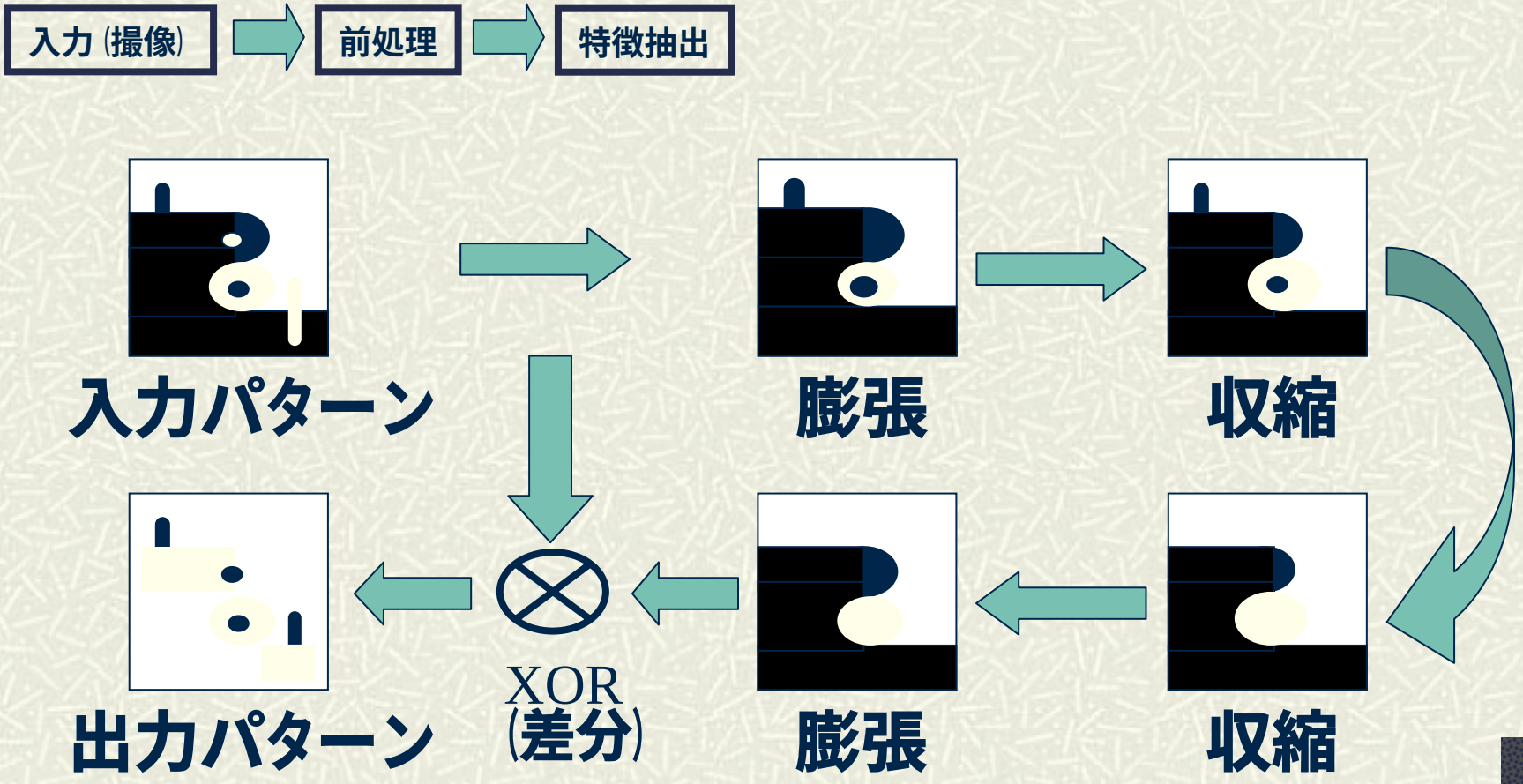
画像認識の処理過程の例
ー　画像強調とシルエットの抽出　ー

入力（撮像） → 前処理 → 特徴抽出

16

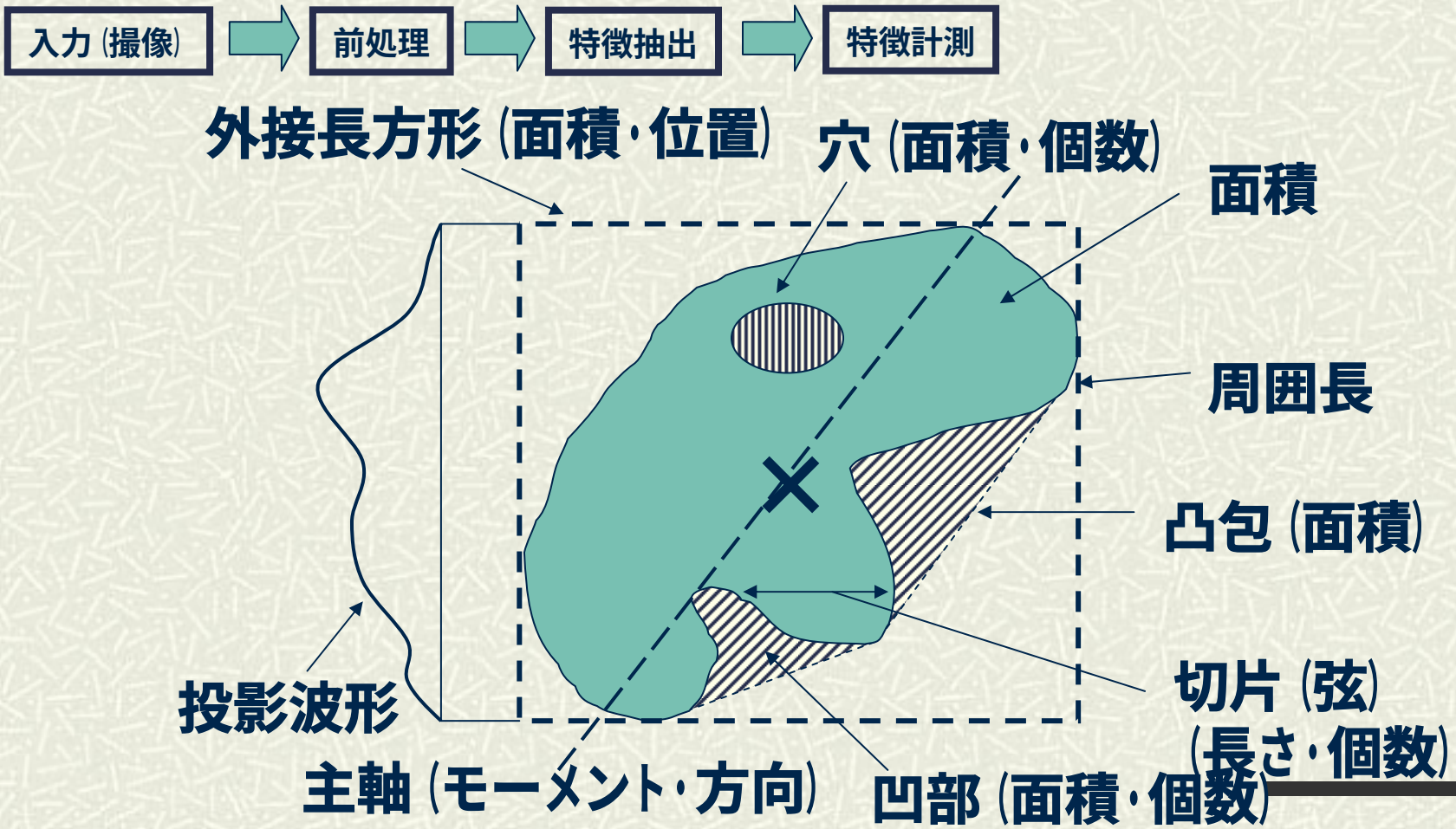
画像認識の処理過程の例

ー 傷（孤立領域・くぼみ・突起）の検出 ー



17

画像認識の処理過程の例
ー　形状特徴の数量化　ー

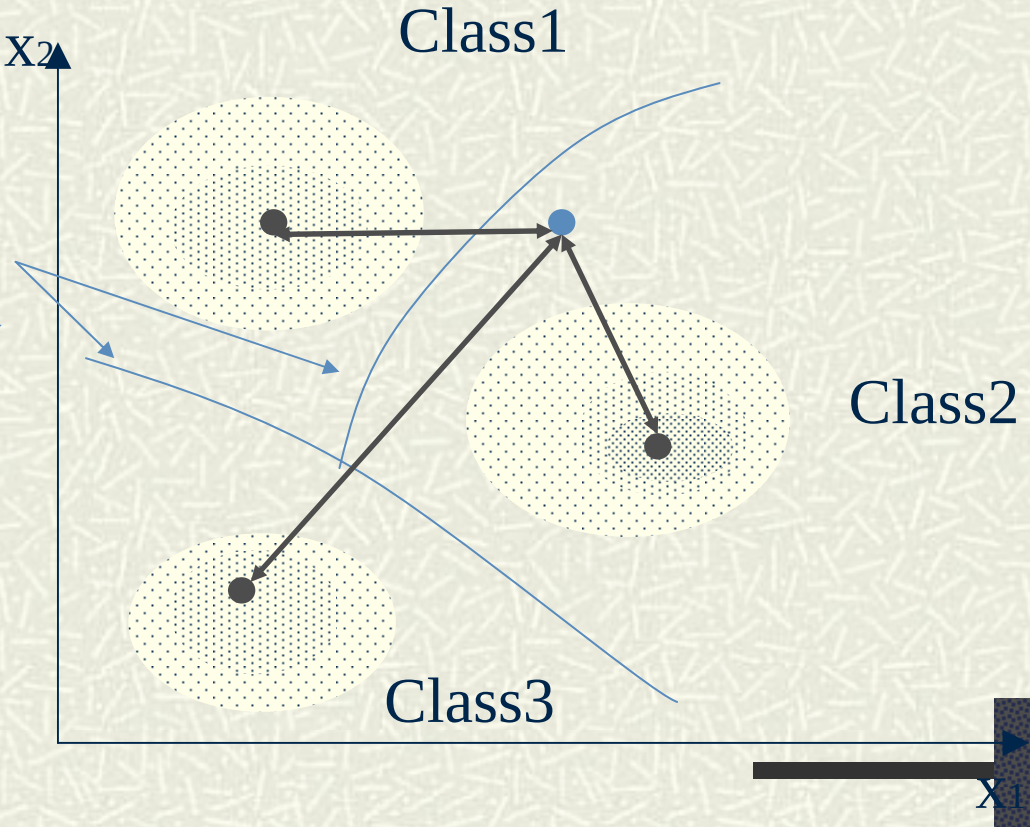


18

画像認識の処理過程の例
ー パターン認識 ー

特徴ベクトル

$$x = \begin{pmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ \cdot \\ X_n \end{pmatrix}$$



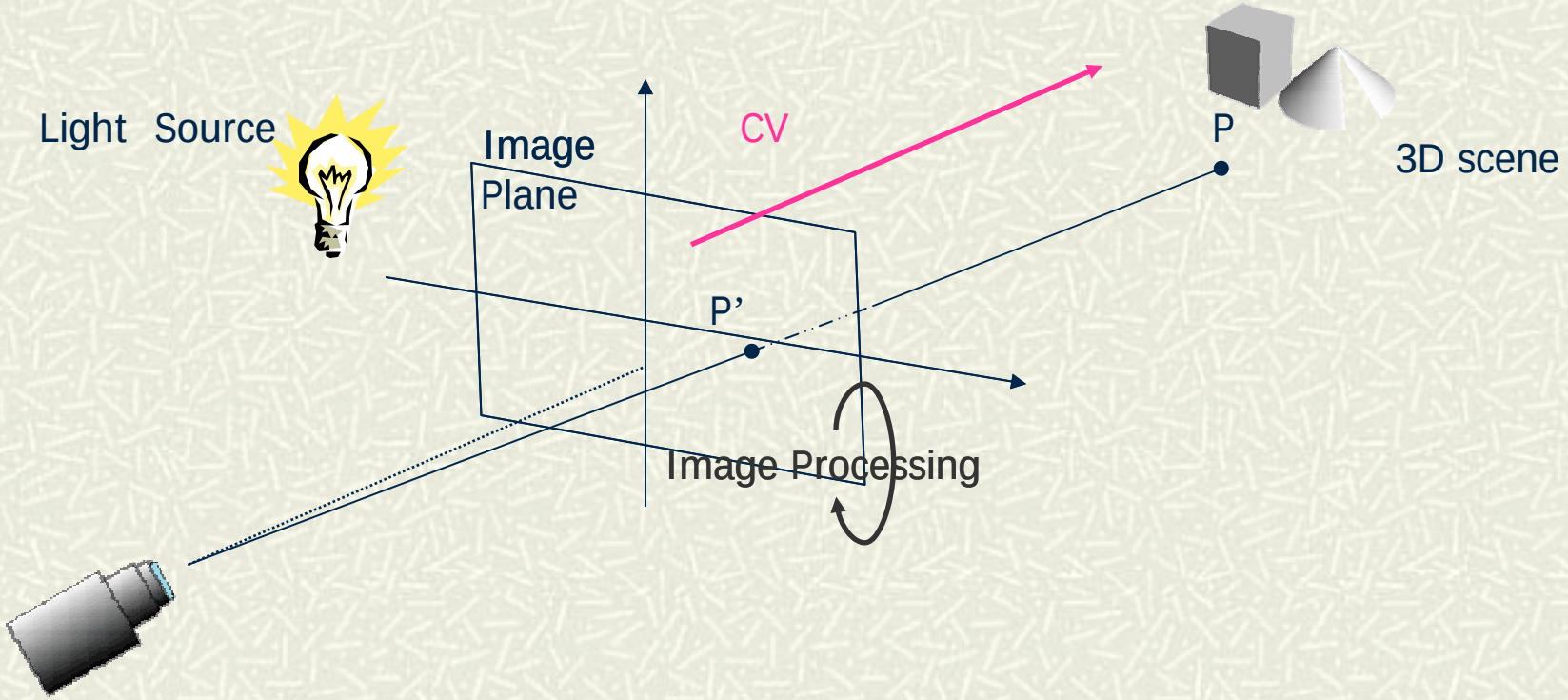
19

コンピュータ・ビジョン

画像処理

↓

コンピュータビジョン



3次元距離計測法

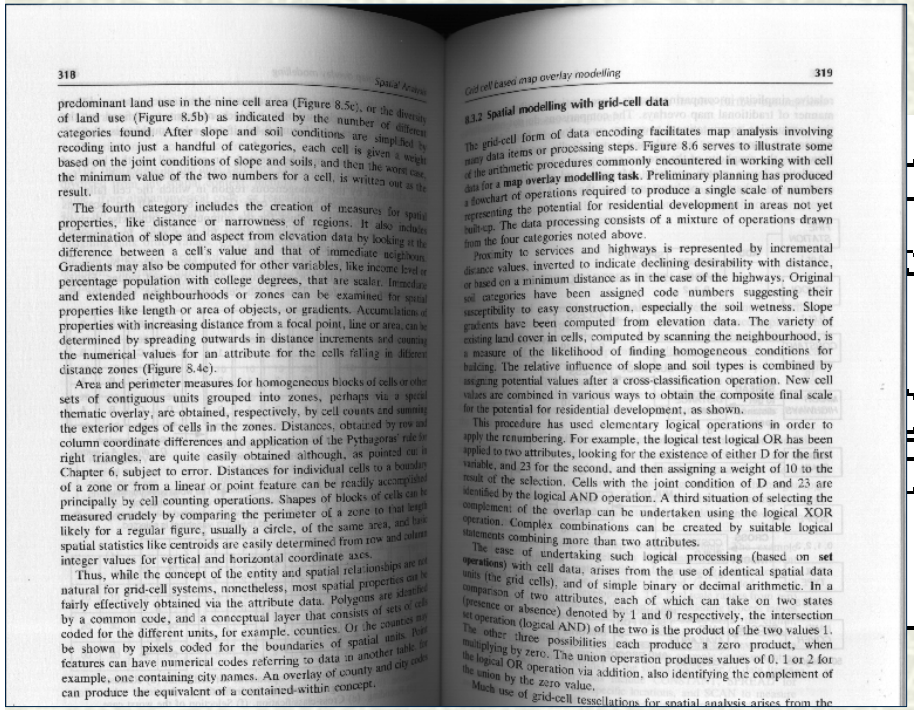
受動型(passive)

- レンズ焦点法
- 単眼視（テクスチャ,影,…）
- ステレオ法
  - 両眼視
  - 三眼視・多眼視
  - カメラ移動型
- 動画像（物体移動）

能動型(active)

- 光レーダー法
  - 位相差計測
  - 時間差計測
- アクティブステレオ法
  - スポット光
  - スリット光・マルチスリット光
  - ステップ光
  - コード化パターン光
  - 濃度・カラー傾斜パターン光
  - 繰り返しパターン光
- 照度差ステレオ法
- モアレ法
- 干渉法
  - ヘテロダイン法
  - タルボー法

コンピュータビジョンの威力 (I)
ー 書籍の高精度コピー ー



318 Spatial Analysis

predominant land use in the nine cell area (Figure 8.5c), or the diversity of land use (Figure 8.5b) as indicated by the number of different categories found. After slope and soil conditions are simplified by recoding into just a handful of categories, each cell is given a weight based on the joint conditions of slope and soils, and then the worst case, the minimum value of the two numbers for a cell, is written out as the result.

The fourth category includes the creation of measures for spatial properties, like distance or narrowness of regions. It also includes determination of slope and aspect from elevation data by looking at the difference between a cell's value and that of immediate neighbours. Gradients may also be computed for other variables, like income level or percentage population with college degrees, that are scalar. Immediate and extended neighbourhoods or zones can be examined for spatial properties like length or area of objects, or gradients. Accumulations of properties with increasing distance from a focal point, line or area, can be determined by spreading outwards in distance increments and counting the numerical values for an attribute for the cells falling in different distance zones (Figure 8.4e).

Area and perimeter measures for homogeneous blocks of cells or other sets of contiguous units grouped into zones, perhaps via a special thematic overlay, are obtained, respectively, by cell counts and summing the exterior edges of cells in the zones. Distances, obtained by row and column coordinate differences and application of the Pythagoras' rule for right triangles, are quite easily obtained although, as pointed out in Chapter 6, subject to error. Distances for individual cells to a boundary of a zone or from a linear or point feature can be readily accomplished principally by cell counting operations. Shapes of blocks of cells can be measured crudely by comparing the perimeter of a zone to that length likely for a regular figure, usually a circle, of the same area, and basic spatial statistics like centroids are easily determined from row and column integer values for vertical and horizontal coordinate axes.

Thus, while the concept of the entity and spatial relationships are not natural for grid-cell systems, nonetheless, most spatial properties can be fairly effectively obtained via the attribute data. Polygons are identified by a common code, and a conceptual layer that consists of sets of cells coded for the different units, for example, counties. Or the counties may be shown by pixels coded for the boundaries of spatial units. Point features can have numerical codes referring to data in another table: for example, one containing city names. An overlay of county and city codes can produce the equivalent of a contained-within concept.

Grid-cell based map overlay modelling 319

8.3.2 Spatial modelling with grid-cell data

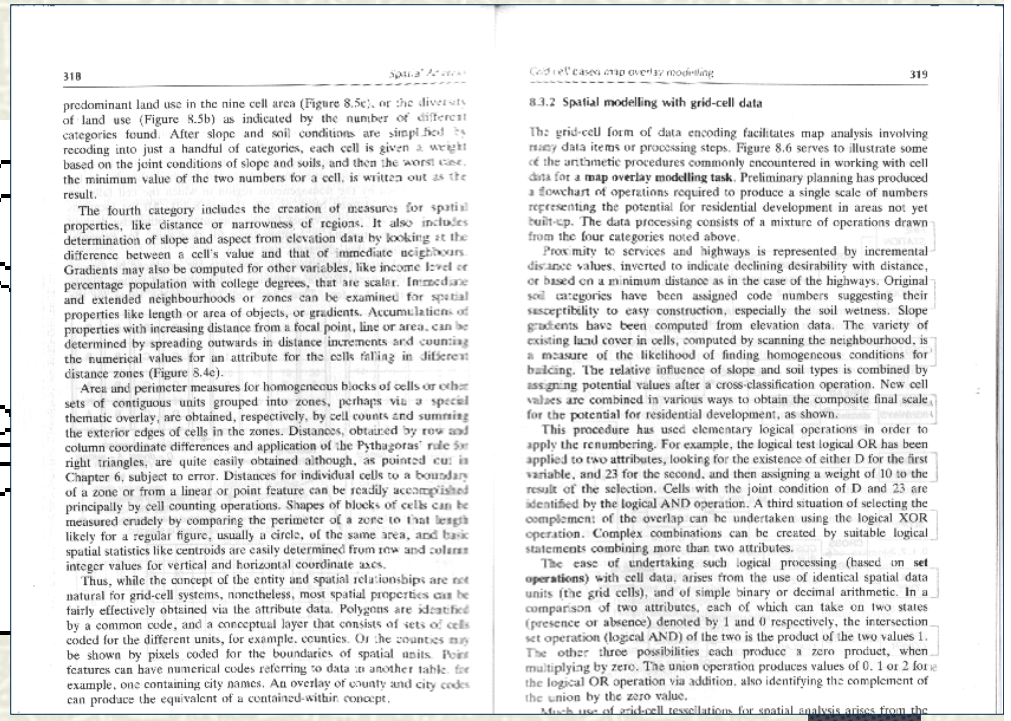
The grid-cell form of data encoding facilitates map analysis involving many data items or processing steps. Figure 8.6 serves to illustrate some of the arithmetic procedures commonly encountered in working with cell data for a **map overlay modelling task**. Preliminary planning has produced a flowchart of operations required to produce a single scale of numbers representing the potential for residential development in areas not yet built-up. The data processing consists of a mixture of operations drawn from the four categories noted above.

Proximity to services and highways is represented by incremental distance values, inverted to indicate declining desirability with distance, or based on a minimum distance as in the case of the highways. Original soil categories have been assigned code numbers suggesting their susceptibility to easy construction, especially the soil wetness. Slope gradients have been computed from elevation data. The variety of existing land cover in cells, computed by scanning the neighbourhood, is a measure of the likelihood of finding homogeneous conditions for building. The relative influence of slope and soil types is combined by assigning potential values after a cross-classification operation. New cell values are combined in various ways to obtain the composite final scale for the potential for residential development, as shown.

This procedure has used elementary logical operations in order to apply the renumbering. For example, the logical test logical OR has been applied to two attributes, looking for the existence of either D for the first variable, and 23 for the second, and then assigning a weight of 10 to the result of the selection. Cells with the joint condition of D and 23 are identified by the logical AND operation. A third situation of selecting the complement of the overlap can be undertaken using the logical XOR operation. Complex combinations can be created by suitable logical statements combining more than two attributes.

The ease of undertaking such logical processing (based on **set operations**) with cell data, arises from the use of identical spatial data units (the grid cells), and of simple binary or decimal arithmetic. In a comparison of two attributes, each of which can take on two states (presence or absence) denoted by 1 and 0 respectively, the intersection set operation (logical AND) of the two is the product of the two values 1. The other three possibilities each produce a zero product, when multiplying by zero. The union operation produces values of 0, 1 or 2 for the logical OR operation via addition, also identifying the complement of the union by the zero value.

Much use of grid-cell tessellations for spatial analysis arises from the



318 Spatial Analysis

predominant land use in the nine cell area (Figure 8.5c), or the diversity of land use (Figure 8.5b) as indicated by the number of different categories found. After slope and soil conditions are simplified by recoding into just a handful of categories, each cell is given a weight based on the joint conditions of slope and soils, and then the worst case, the minimum value of the two numbers for a cell, is written out as the result.

The fourth category includes the creation of measures for spatial properties, like distance or narrowness of regions. It also includes determination of slope and aspect from elevation data by looking at the difference between a cell's value and that of immediate neighbours. Gradients may also be computed for other variables, like income level or percentage population with college degrees, that are scalar. Immediate and extended neighbourhoods or zones can be examined for spatial properties like length or area of objects, or gradients. Accumulations of properties with increasing distance from a focal point, line or area, can be determined by spreading outwards in distance increments and counting the numerical values for an attribute for the cells falling in different distance zones (Figure 8.4e).

Area and perimeter measures for homogeneous blocks of cells or other sets of contiguous units grouped into zones, perhaps via a special thematic overlay, are obtained, respectively, by cell counts and summing the exterior edges of cells in the zones. Distances, obtained by row and column coordinate differences and application of the Pythagoras' rule for right triangles, are quite easily obtained although, as pointed out in Chapter 6, subject to error. Distances for individual cells to a boundary of a zone or from a linear or point feature can be readily accomplished principally by cell counting operations. Shapes of blocks of cells can be measured crudely by comparing the perimeter of a zone to that length likely for a regular figure, usually a circle, of the same area, and basic spatial statistics like centroids are easily determined from row and column integer values for vertical and horizontal coordinate axes.

Thus, while the concept of the entity and spatial relationships are not natural for grid-cell systems, nonetheless, most spatial properties can be fairly effectively obtained via the attribute data. Polygons are identified by a common code, and a conceptual layer that consists of sets of cells coded for the different units, for example, counties. Or the counties may be shown by pixels coded for the boundaries of spatial units. Point features can have numerical codes referring to data in another table: for example, one containing city names. An overlay of county and city codes can produce the equivalent of a contained-within concept.

Grid-cell based map overlay modelling 319

8.3.2 Spatial modelling with grid-cell data

The grid-cell form of data encoding facilitates map analysis involving many data items or processing steps. Figure 8.6 serves to illustrate some of the arithmetic procedures commonly encountered in working with cell data for a **map overlay modelling task**. Preliminary planning has produced a flowchart of operations required to produce a single scale of numbers representing the potential for residential development in areas not yet built-up. The data processing consists of a mixture of operations drawn from the four categories noted above.

Proximity to services and highways is represented by incremental distance values, inverted to indicate declining desirability with distance, or based on a minimum distance as in the case of the highways. Original soil categories have been assigned code numbers suggesting their susceptibility to easy construction, especially the soil wetness. Slope gradients have been computed from elevation data. The variety of existing land cover in cells, computed by scanning the neighbourhood, is a measure of the likelihood of finding homogeneous conditions for building. The relative influence of slope and soil types is combined by assigning potential values after a cross-classification operation. New cell values are combined in various ways to obtain the composite final scale for the potential for residential development, as shown.

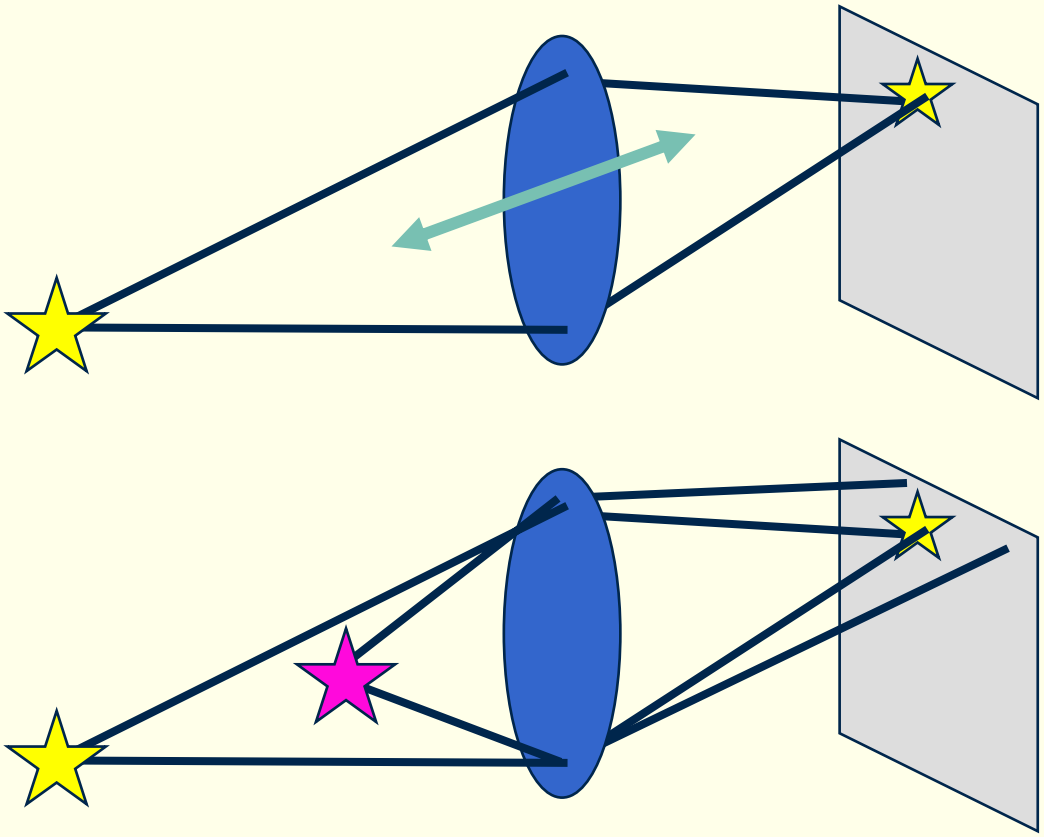
This procedure has used elementary logical operations in order to apply the renumbering. For example, the logical test logical OR has been applied to two attributes, looking for the existence of either D for the first variable, and 23 for the second, and then assigning a weight of 10 to the result of the selection. Cells with the joint condition of D and 23 are identified by the logical AND operation. A third situation of selecting the complement of the overlap can be undertaken using the logical XOR operation. Complex combinations can be created by suitable logical statements combining more than two attributes.

The ease of undertaking such logical processing (based on **set operations**) with cell data, arises from the use of identical spatial data units (the grid cells), and of simple binary or decimal arithmetic. In a comparison of two attributes, each of which can take on two states (presence or absence) denoted by 1 and 0 respectively, the intersection set operation (logical AND) of the two is the product of the two values 1. The other three possibilities each produce a zero product, when multiplying by zero. The union operation produces values of 0, 1 or 2 for the logical OR operation via addition, also identifying the complement of the union by the zero value.

Much use of grid-cell tessellations for spatial analysis arises from the

23

Depth from Defocus



Depth from Focusing

レンズを駆動し合焦点を探索 - 物理的駆動が介在するため実時間計測に向かない

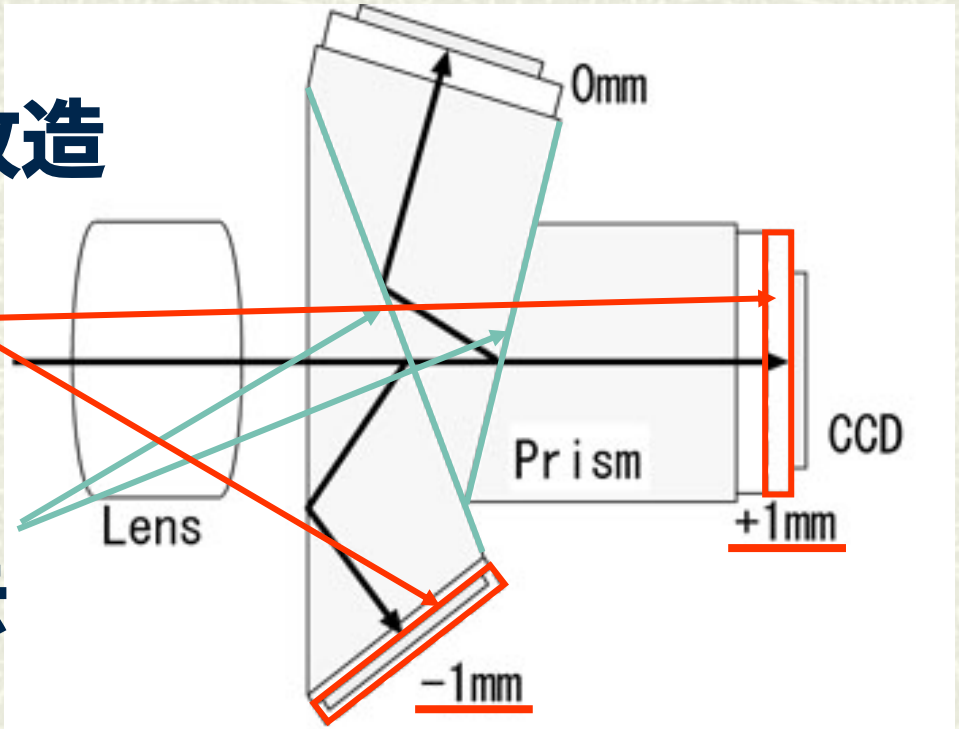
Depth from Defocus

ぼけ量の解析から距離を推定 - 実時間計測が可能機械的駆動部がない

24

多重フォーカスカメラの構造

- カラーCCDカメラを改造
- 各CCDを1mmづつ光軸方向に移動
- プリズムの再蒸着による分光特性の除去
- 通常の3CCDカメラと同等の寸法／重量



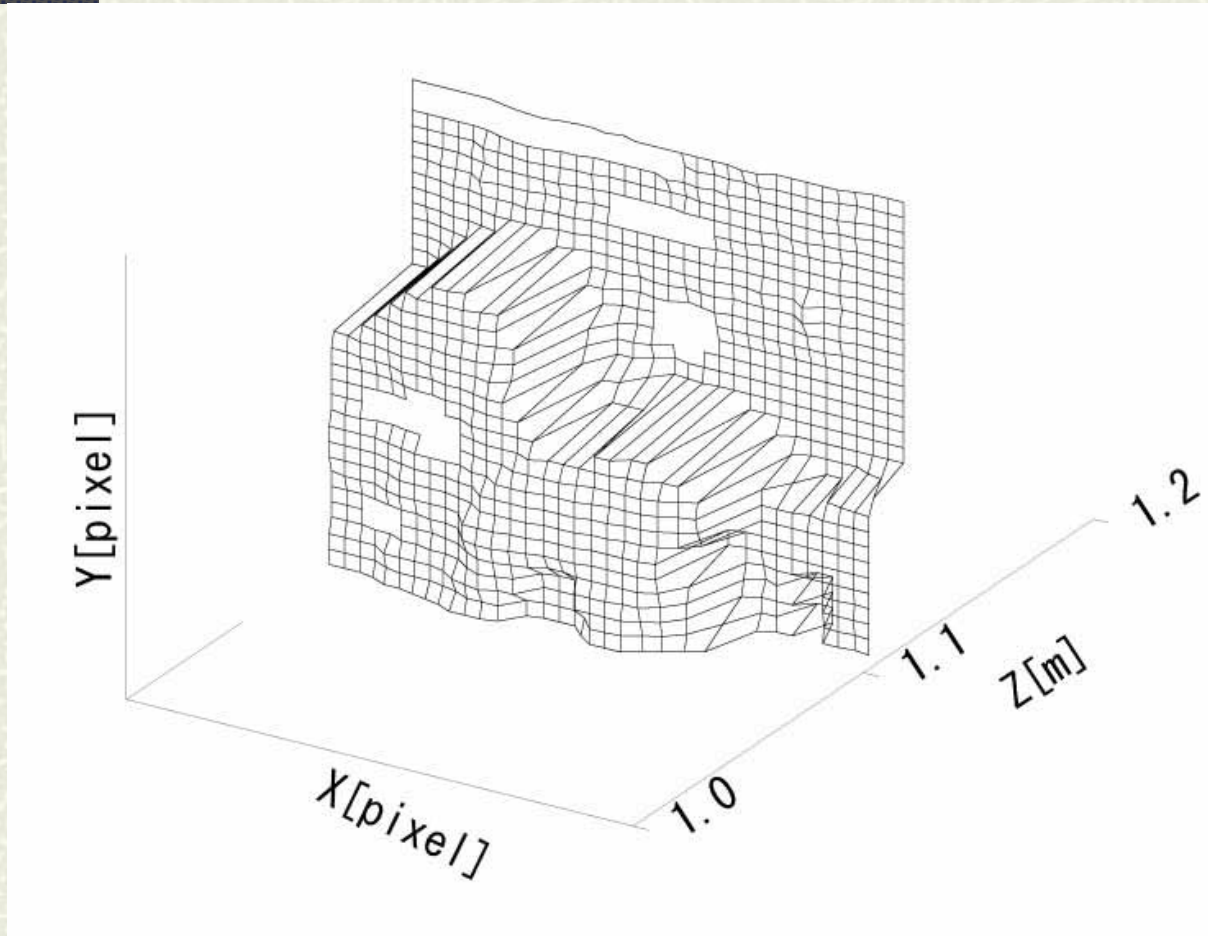
25

実験：入力画像

最後面CCD

26

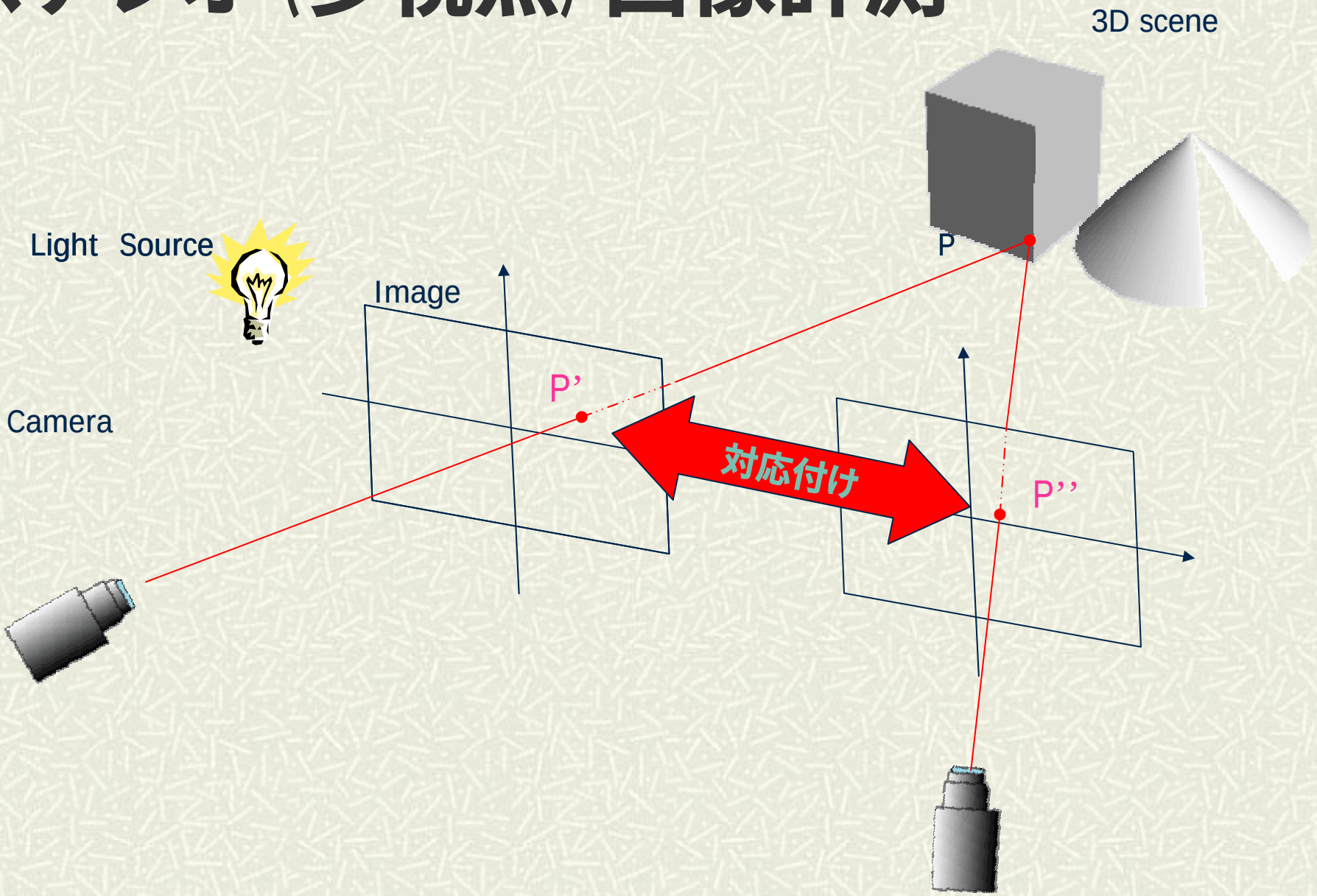
実験：復元された対象の形状



対象の三次元形状

完全合焦画像（部分）

ステレオ (多視点) 画像計測



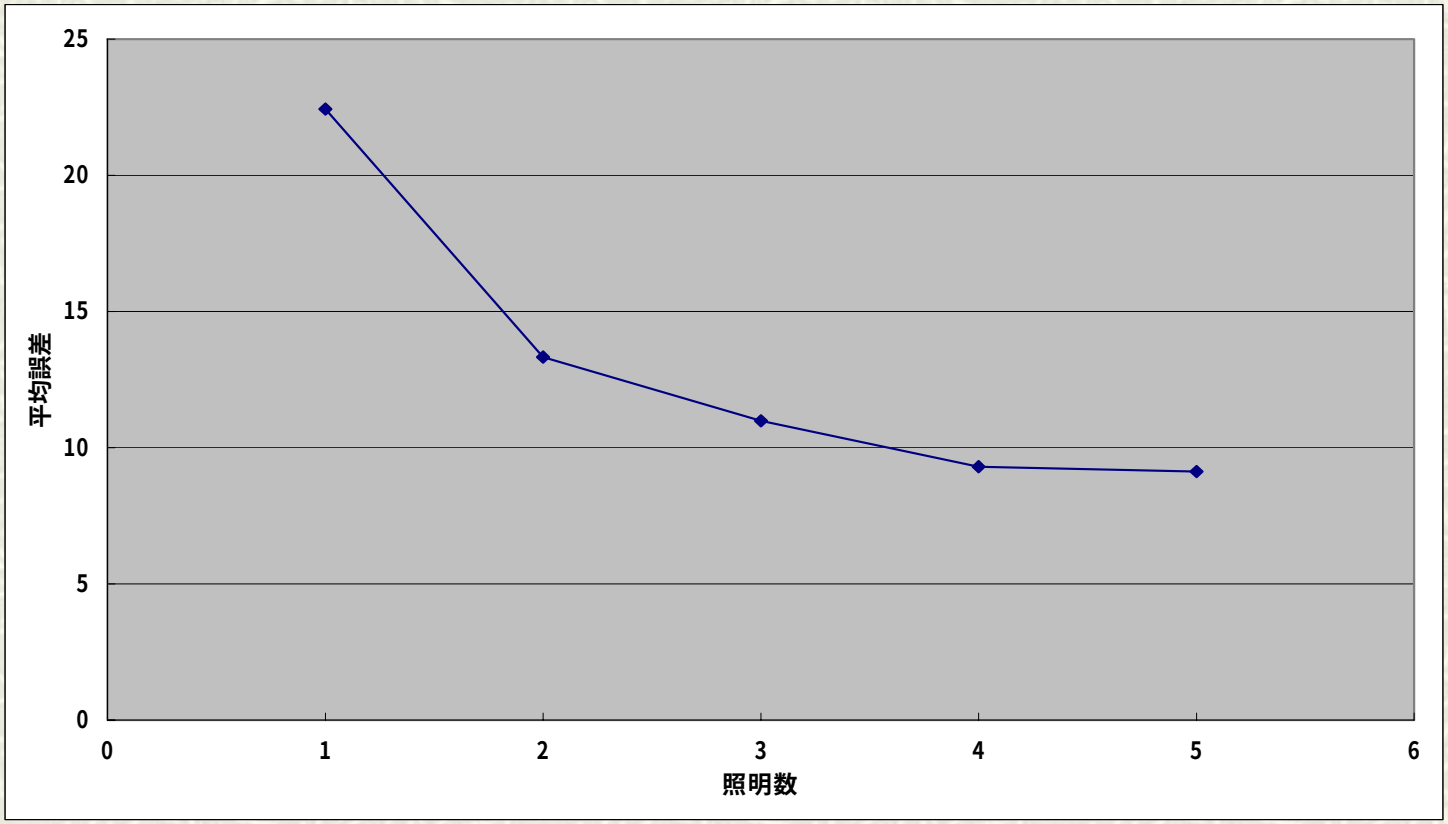
路床の画像

29

多重光源ステレオ画像解析

30

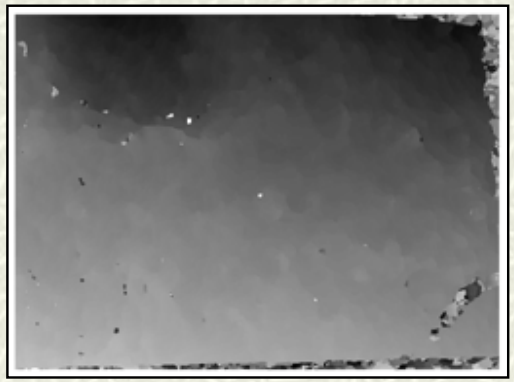
性能評価



31

路床の異常部検出結果

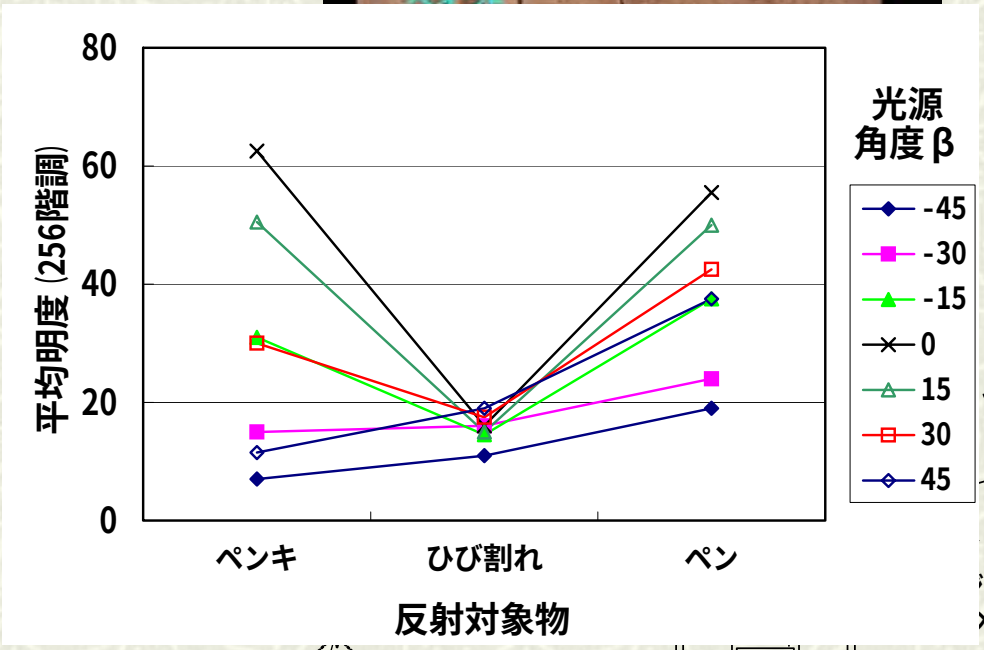
ステレオ画像



3次元距離画像　　検出された異常部

32

複数光源を利用したトンネルのひび割れ検出法



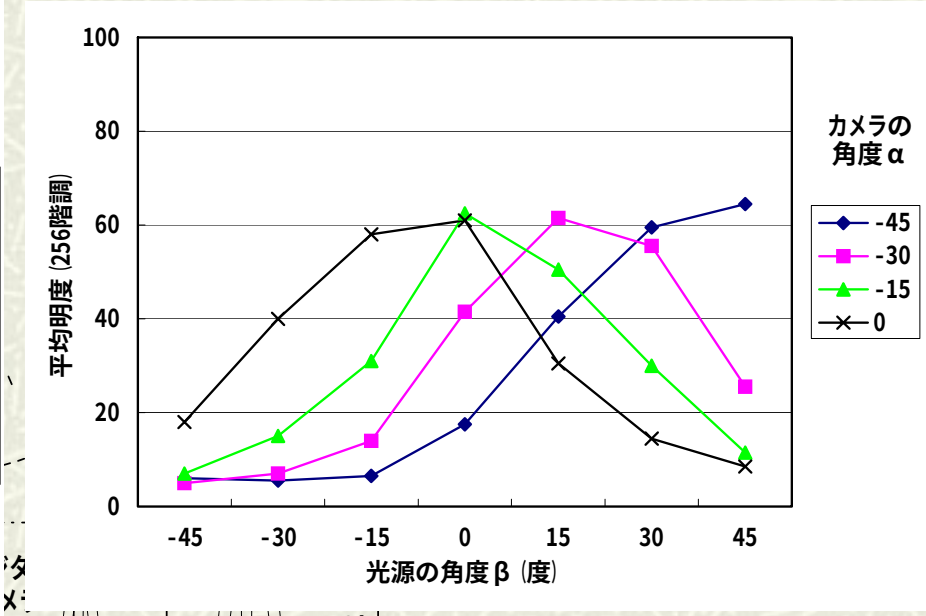
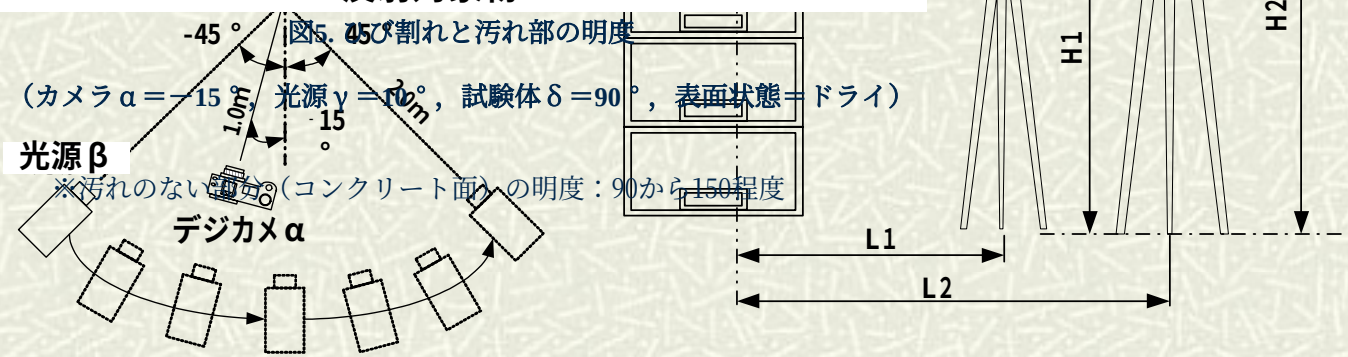


図5. ひび割れと汚れ部の明度

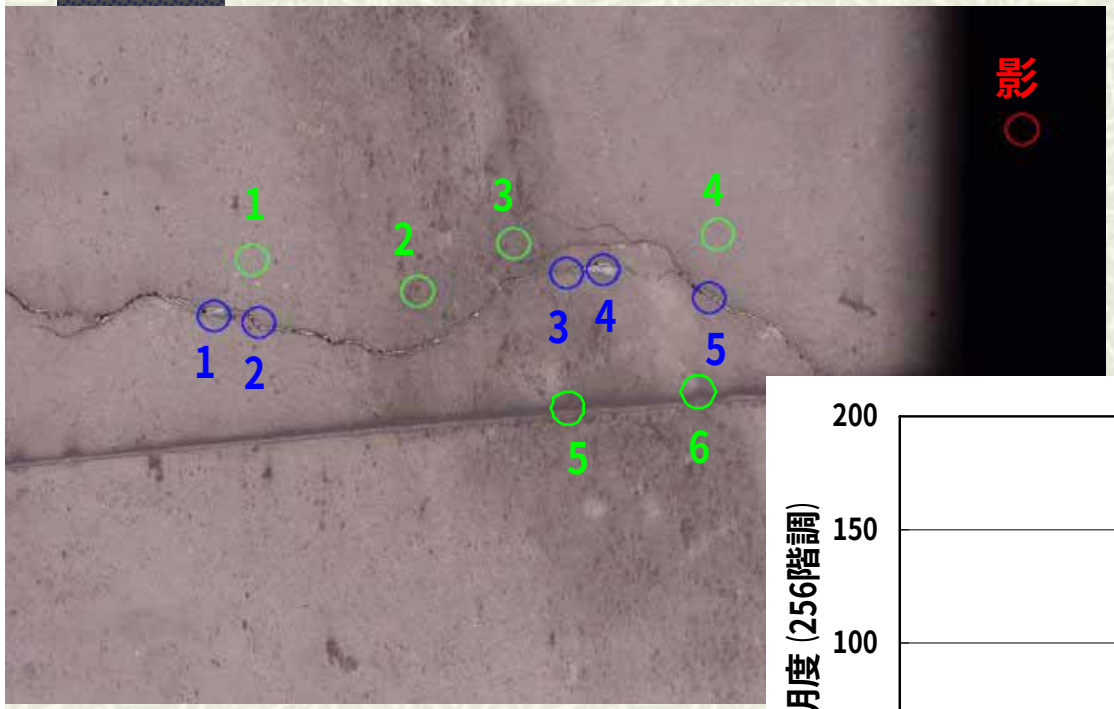
（カメラα＝－15°，光源γ＝10°，試験体δ＝90°，表面状態＝ドライ）

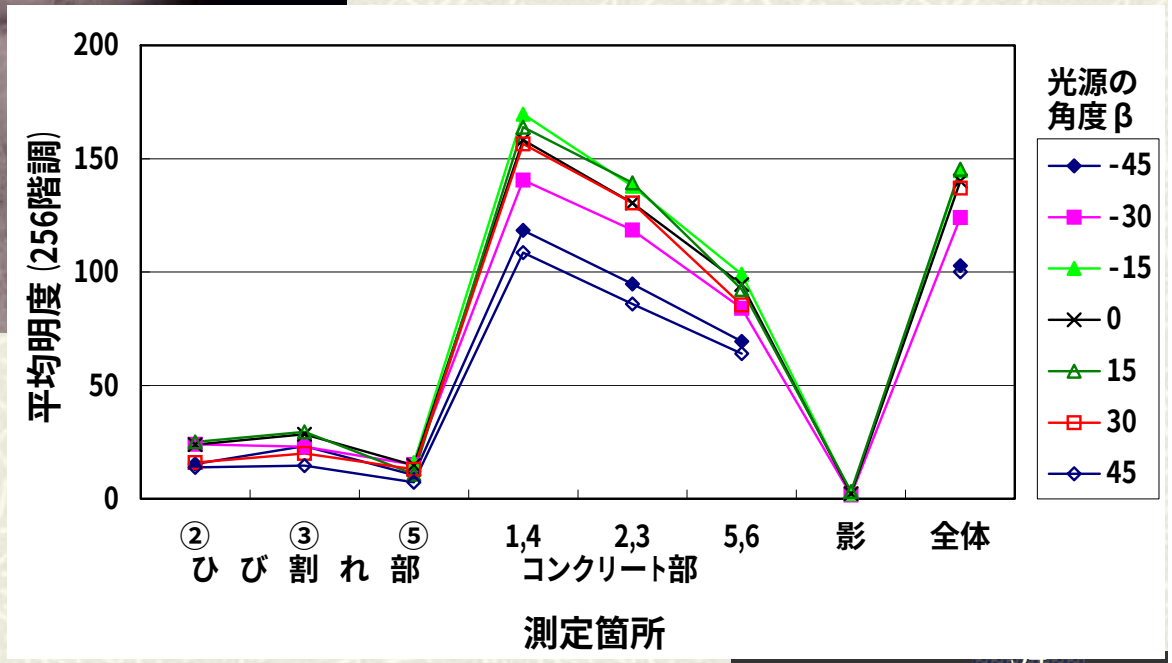


※汚れのない部分（コンクリート面）の明度：90から150程度

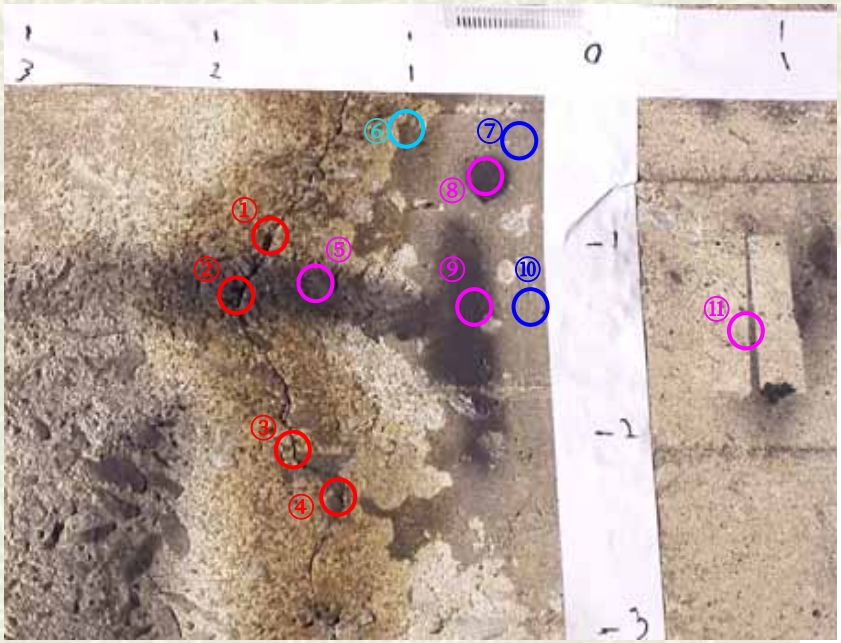
33

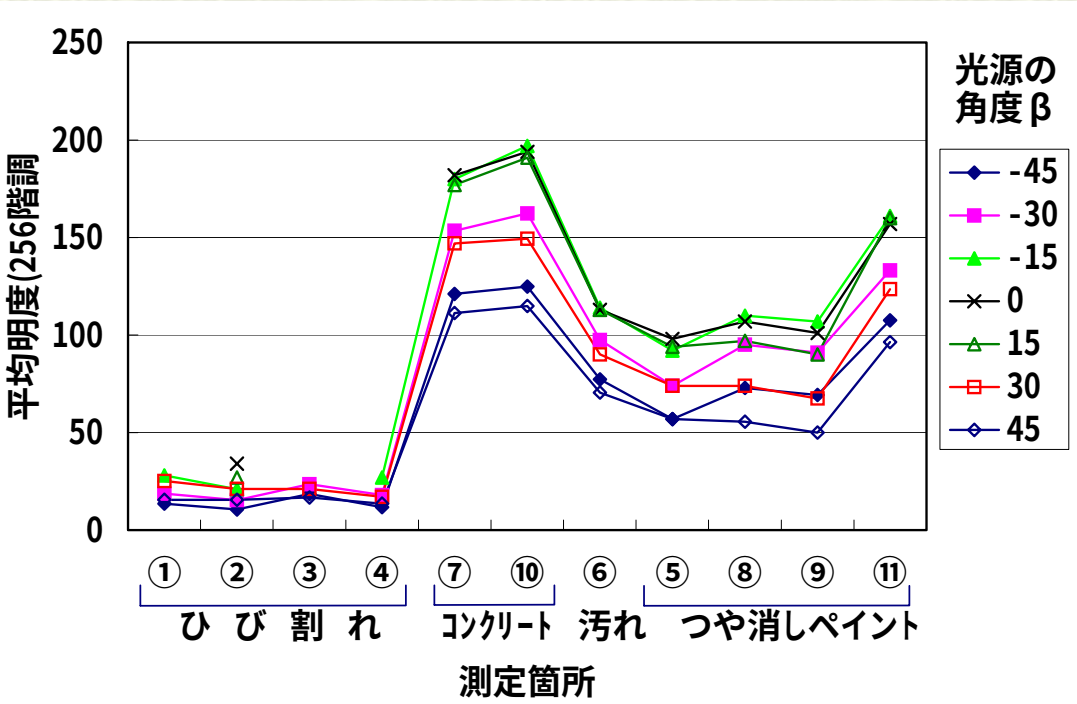
現場での性能評価 (I)


影
1
2
3
4
1
2
3
4
5
5
6


200
150
100
50
0
平均明度 (256階調)
光源の角度β
-45
-30
-15
0
15
30
45
②
③
⑤
1,4
2,3
5,6
影
全体
ひ び 割 れ 部
コンクリート部
測定箇所

現場での性能評価 (II)





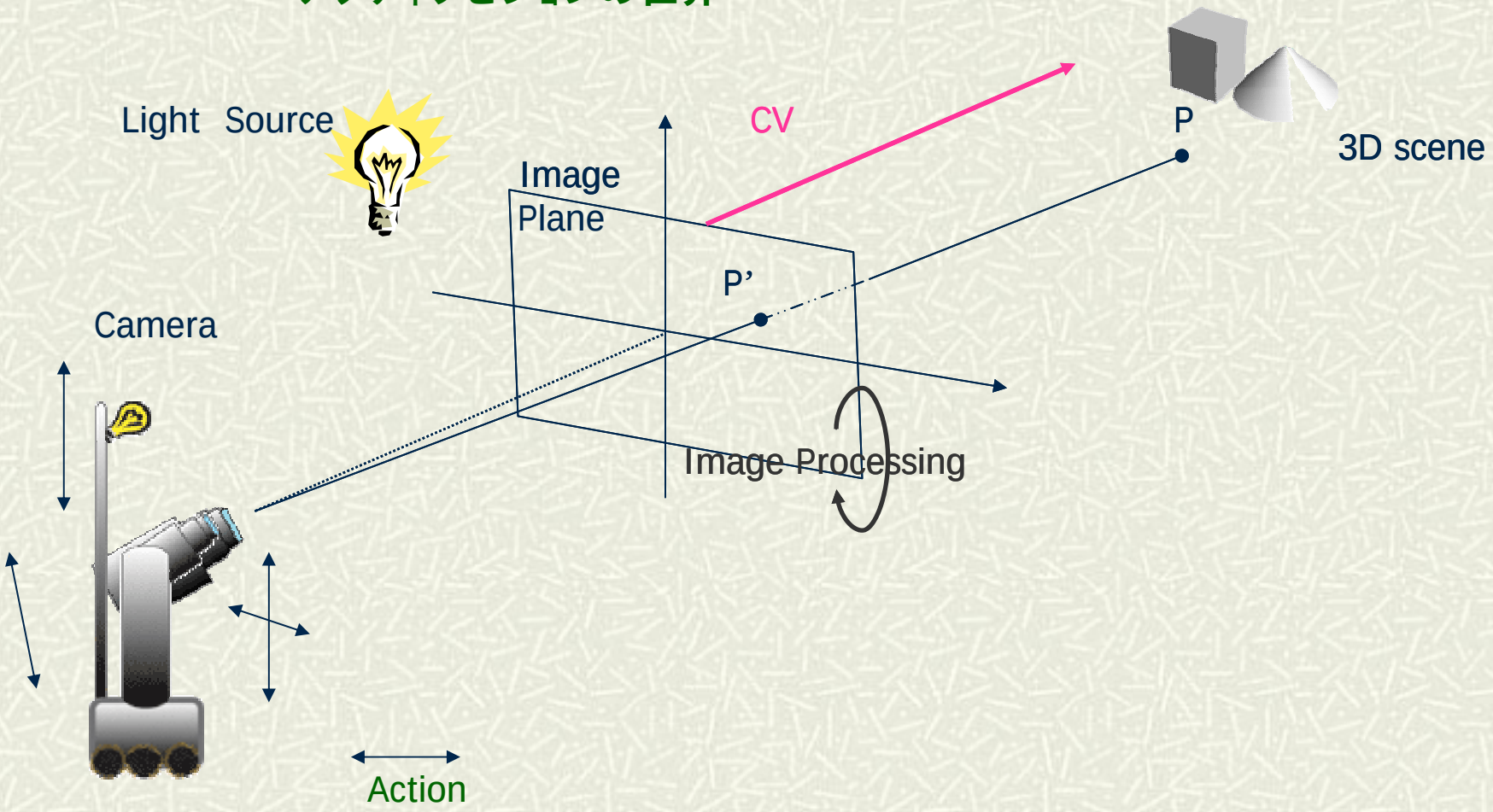
35

アクティブ・ビジョン

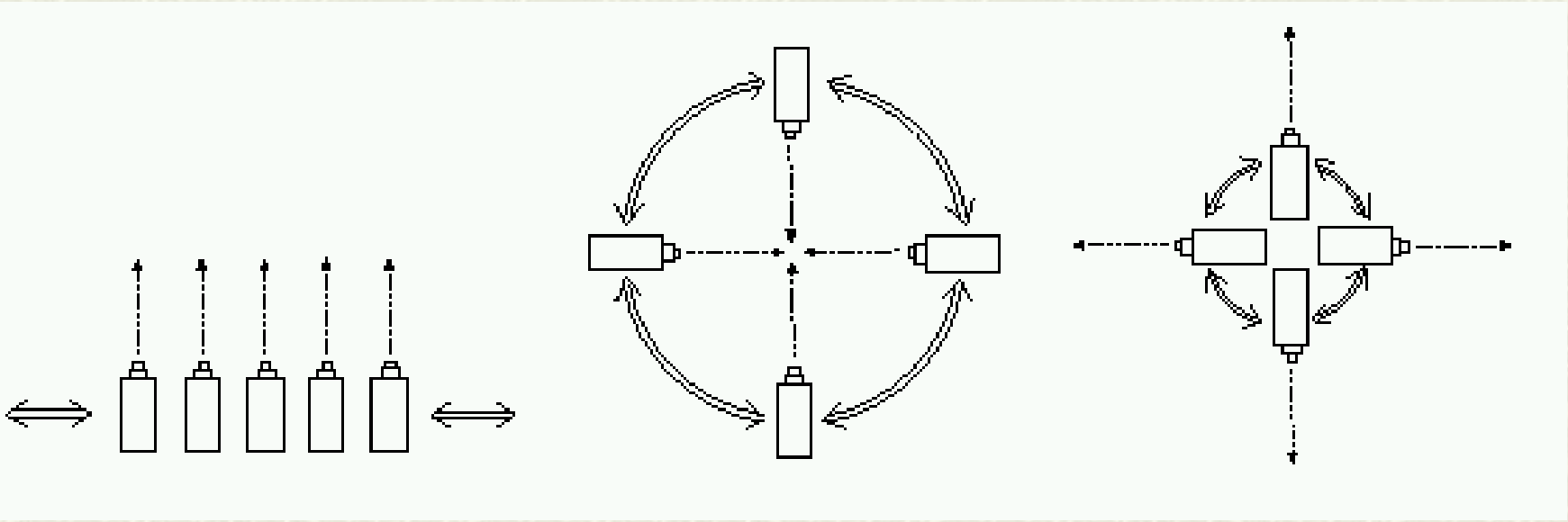
画像処理

コンピュータビジョン

アクティブビジョンの世界



多重カメラ配置による画像計測



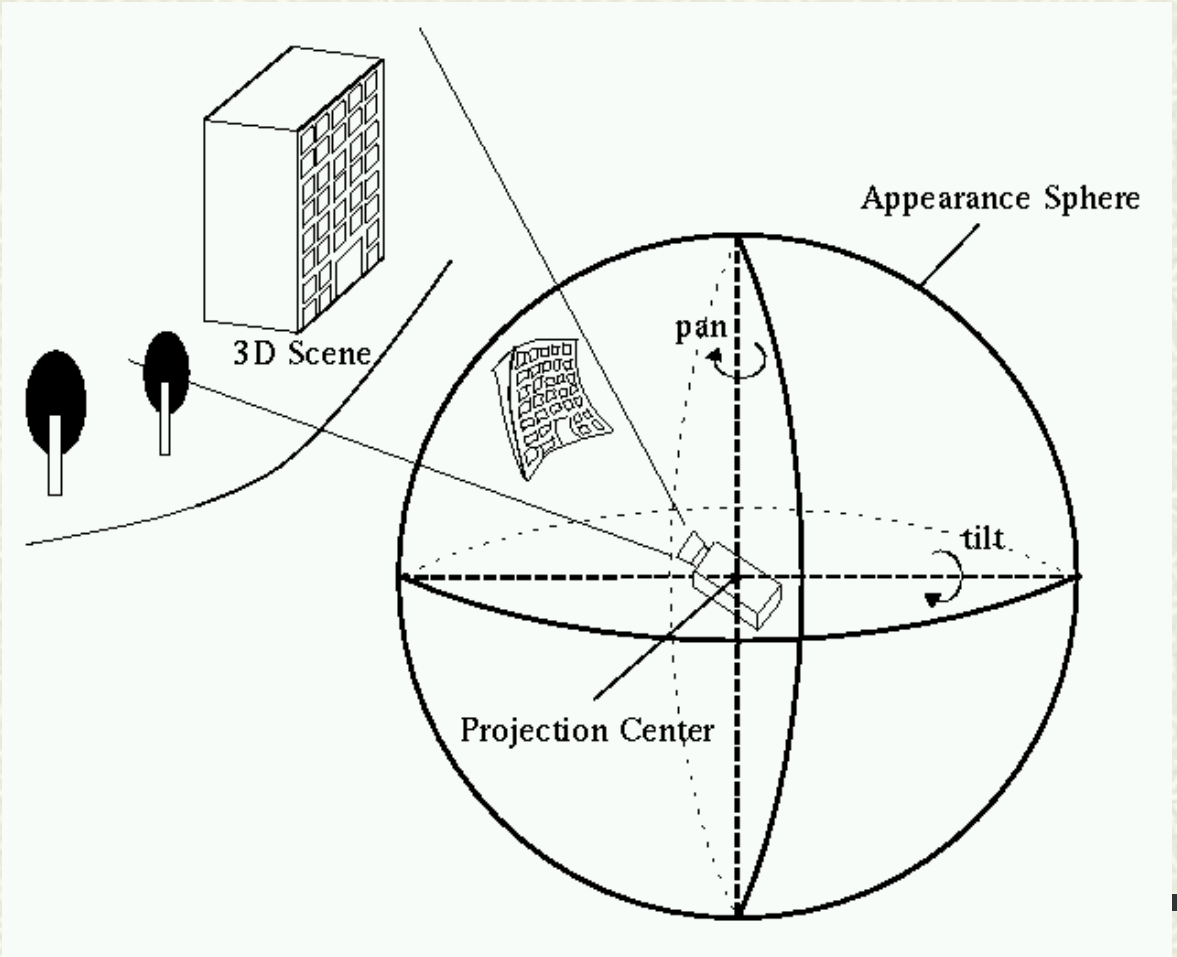
並行配置　　収斂配置　　発散配置

広域観測　　3次元観測　　全方位観測

38

アクティブカメラの威力
ー　視点固定型パン･チルト･ズームカメラ　ー

投影中心　＝　回転中心

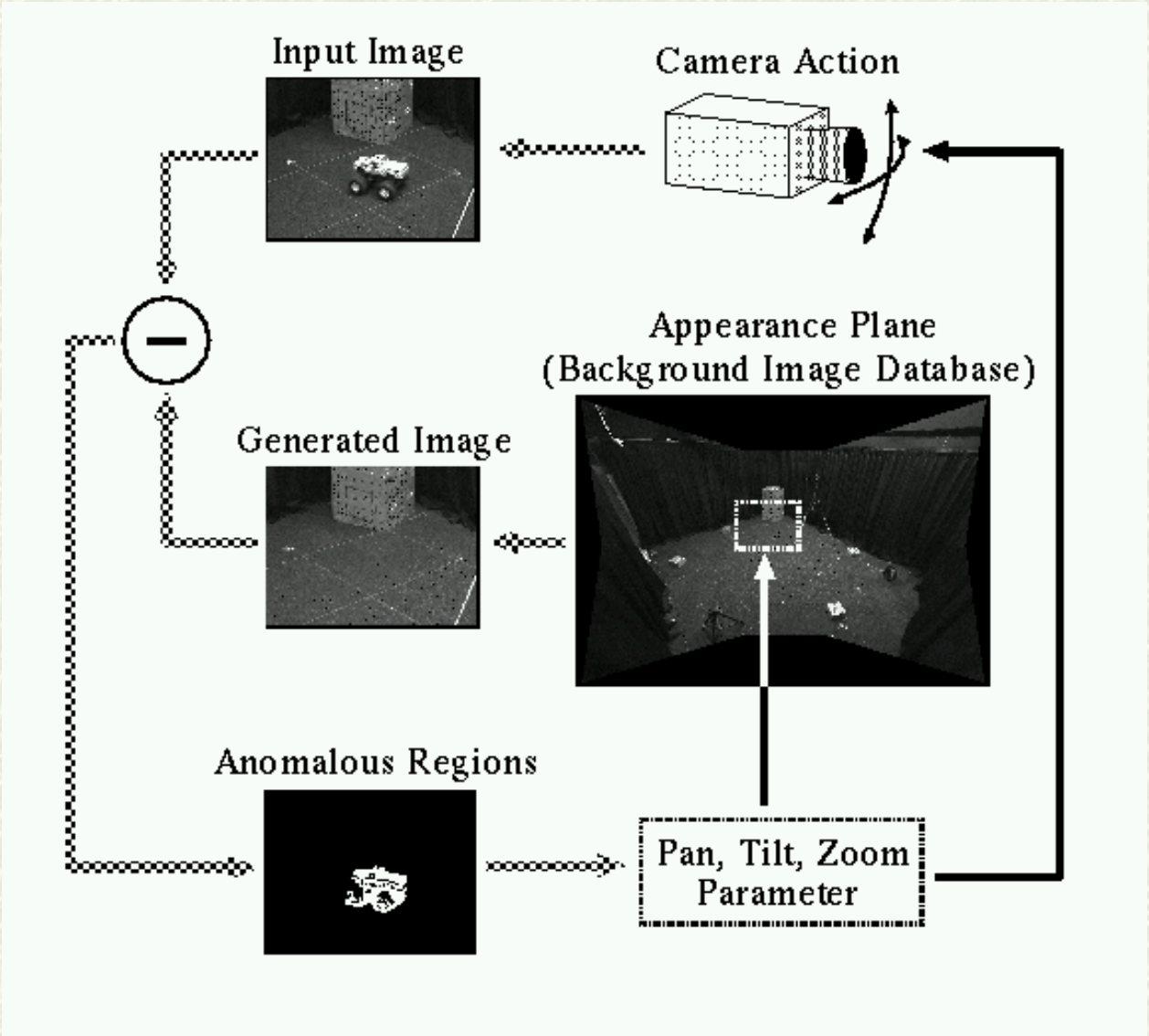


39

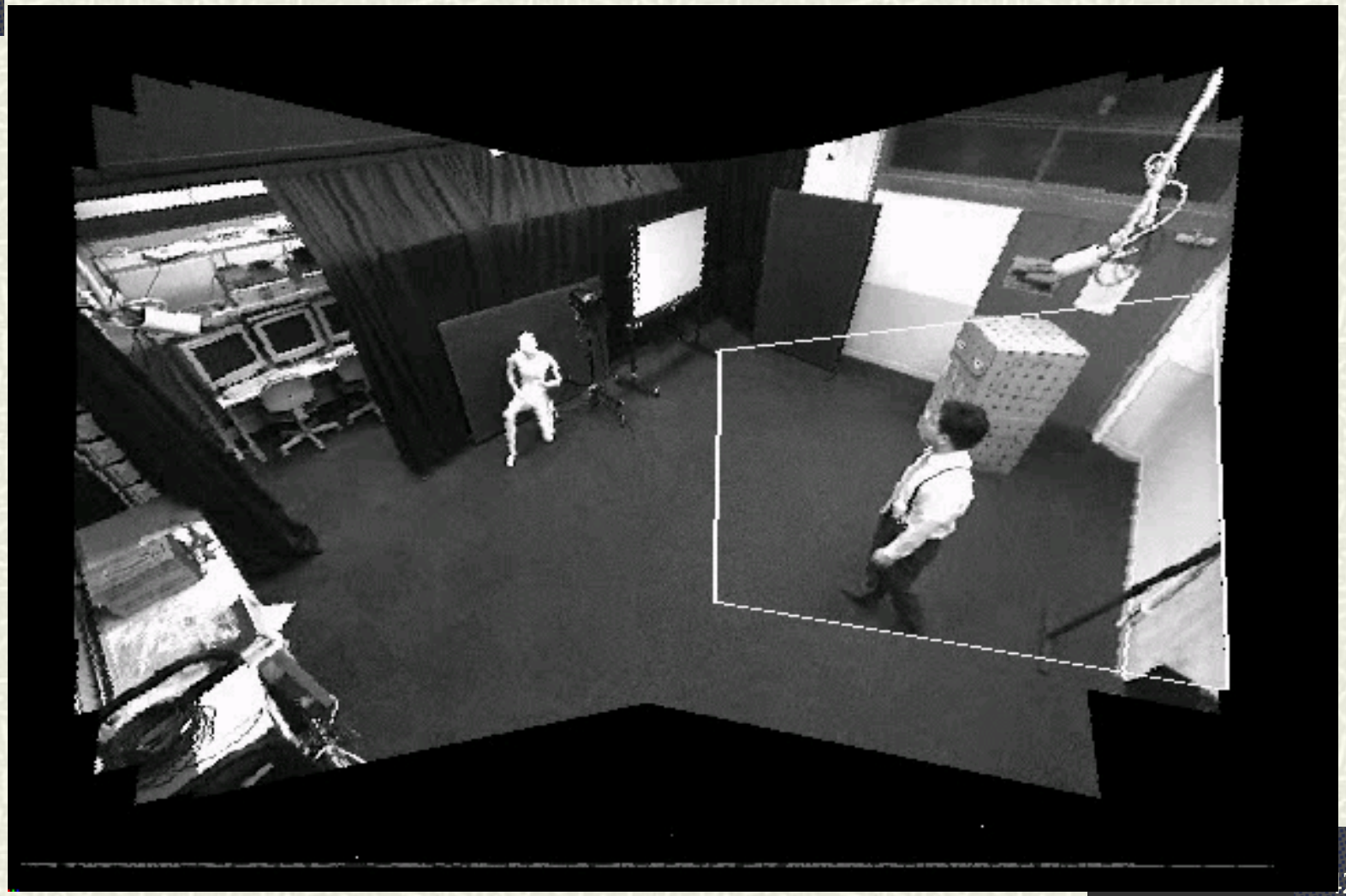
視点固定型パン･チルトカメラ

40

視点固定型パン･チルト･ズームカメラを使った能動的対象追跡アルゴリズム



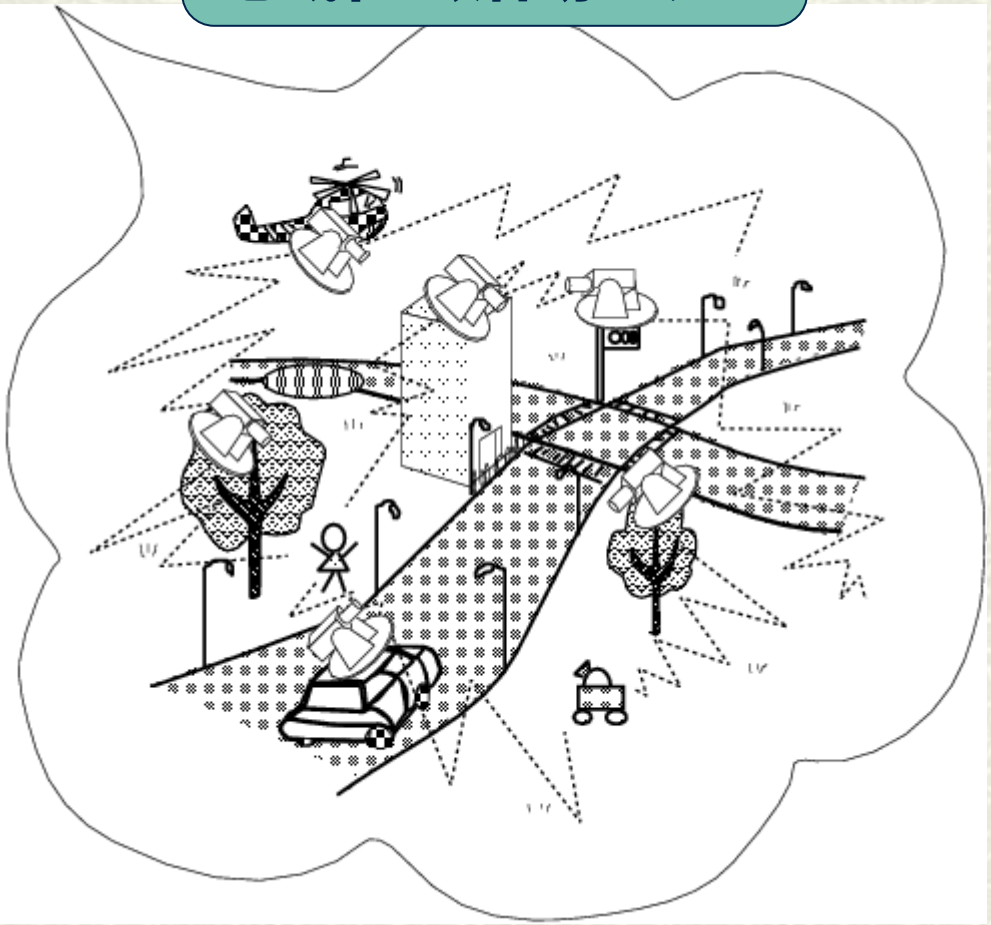
局所的追跡ビデオ映像と
大局的パノラマ画像の統合表示



センサ・ネットワーク
ー ユビキタス社会の実現 ー

センサーネットワークの応用
ー 交通･人の流れ･環境のモニタリング ー

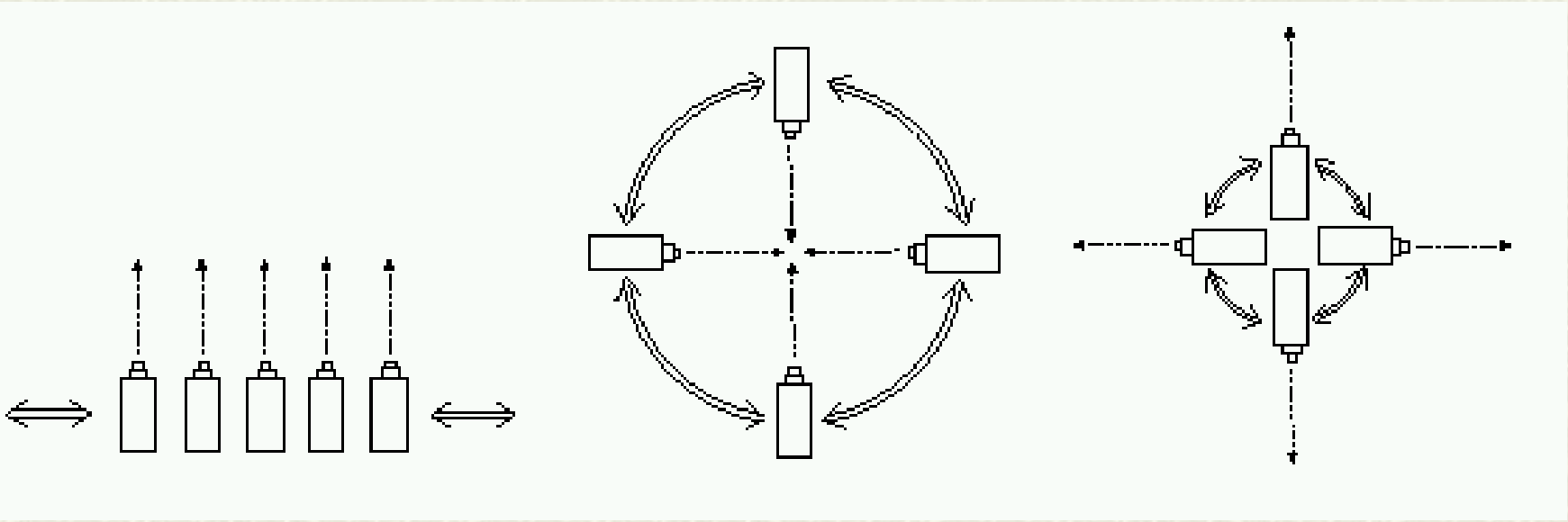
物流･人流に基づく地域社会活動モデル



フィールド･サーバ温度･湿度･日照計測

44

多重カメラ配置による画像計測



並行配置　　収斂配置　　発散配置

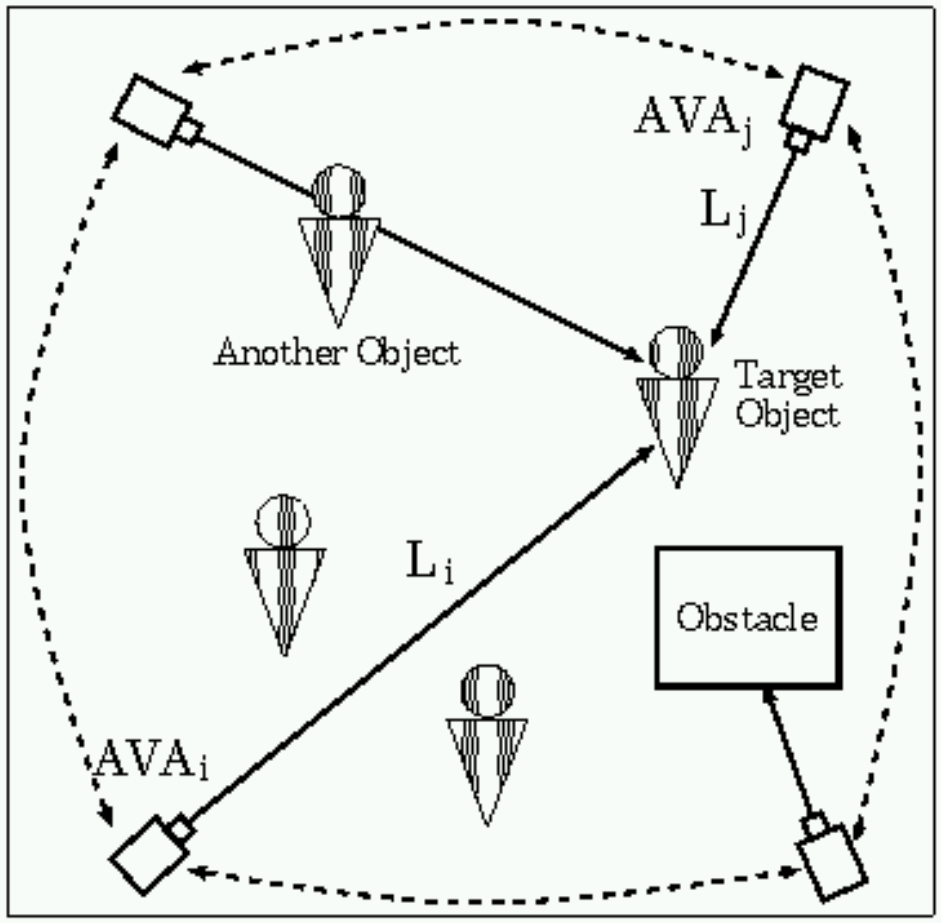
広域観測　　3次元観測　　全方位観測

45

分散配置された能動カメラ群による複数対象の協調追跡

ネットワーク
通信機能

能動視覚エージェント



視覚認識機能

行動（カメラ制御）機能

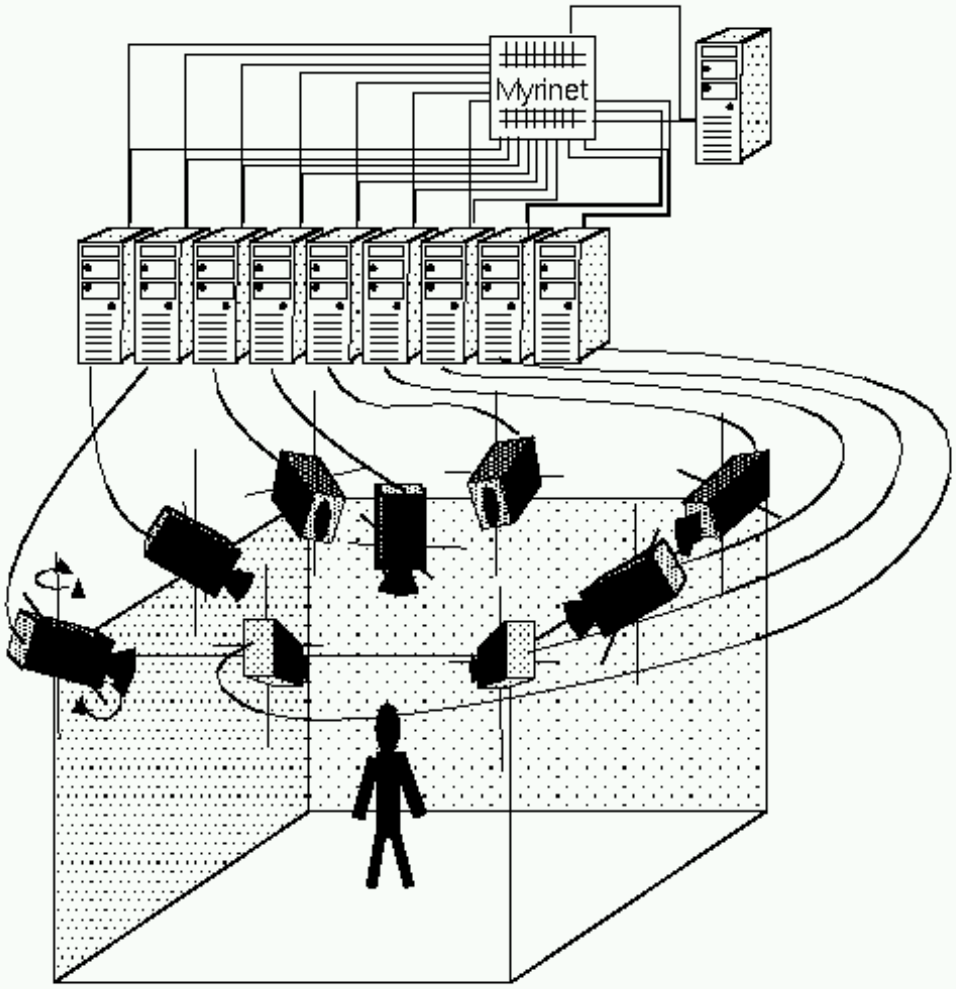
46

能動視覚エージェント群による複数対象の協調追跡

47

3次元ビデオ撮影システム

Myrinet
(1.28Gbps)
PC Cluster



ネットワークソフト：
PM Library
(RWC 開発)

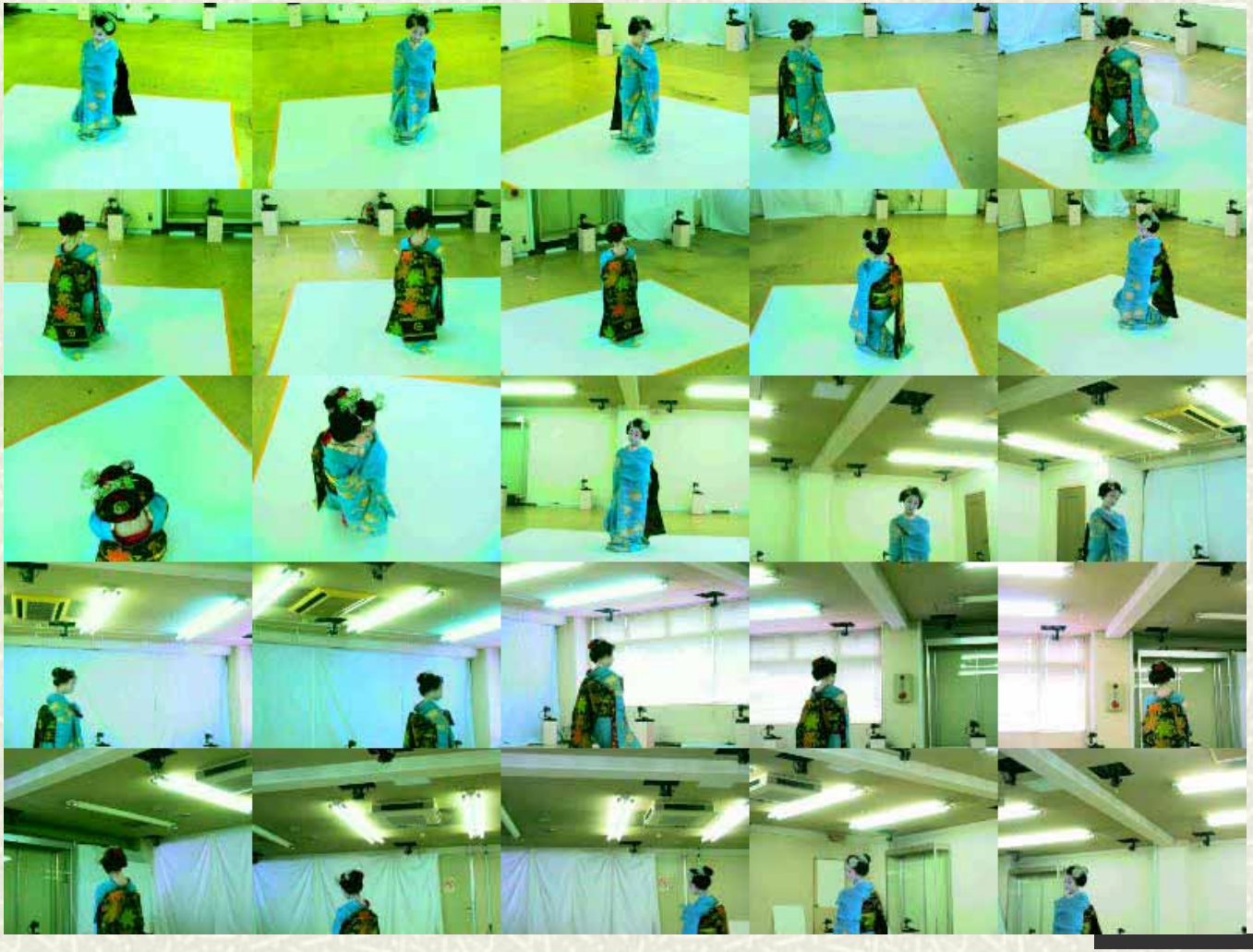
PC: 30台

カメラ：25台

視点固定型
パン・チルト・ズーム
カメラ

48

多数の視点から同時撮影されたビデオ



49

3次元ビデオ生成の処理過程

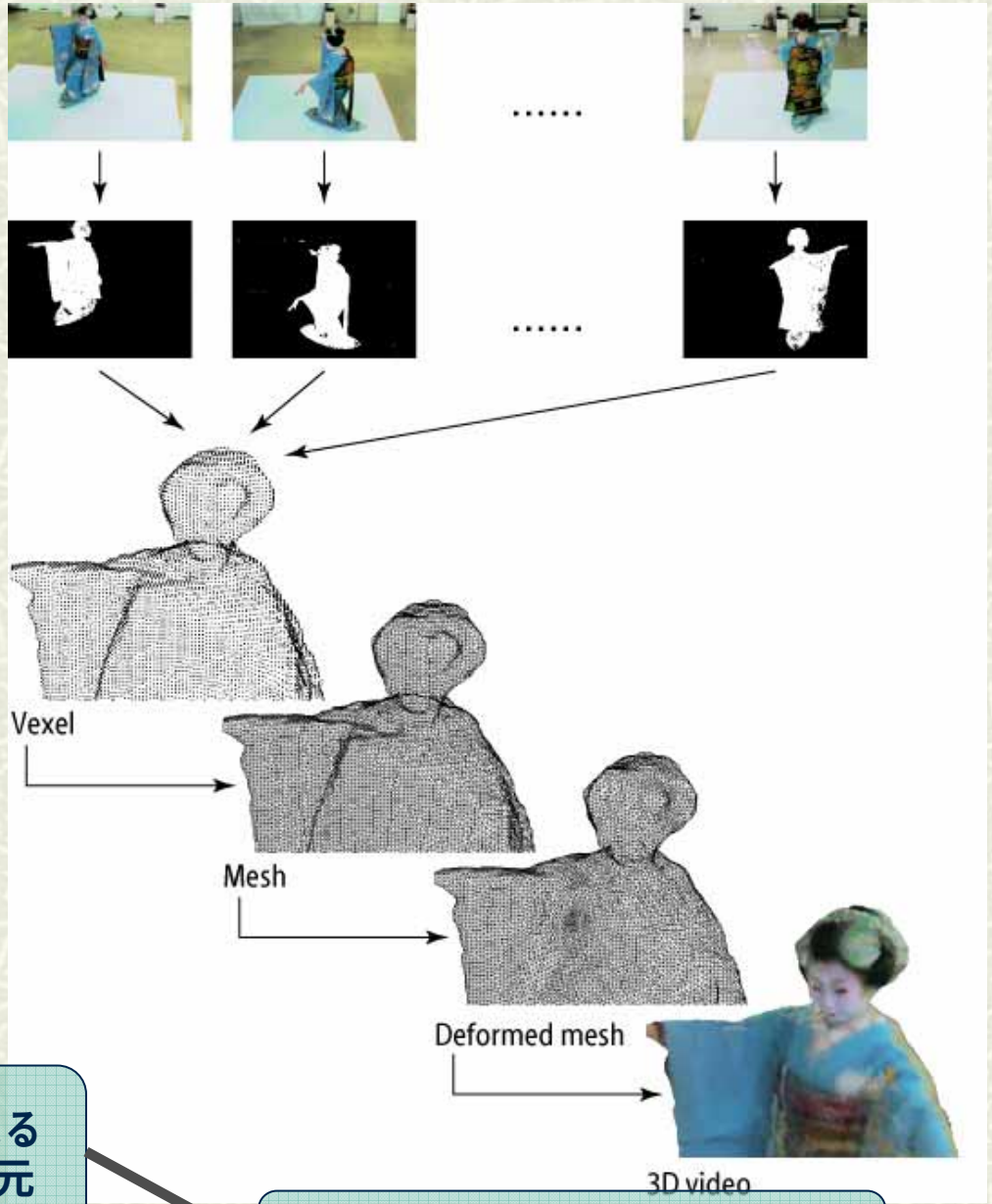
カメラキャリブレーション

多視点ビデオの同期撮影

対象シルエット抽出

視体積交差法による
3次元形状復元

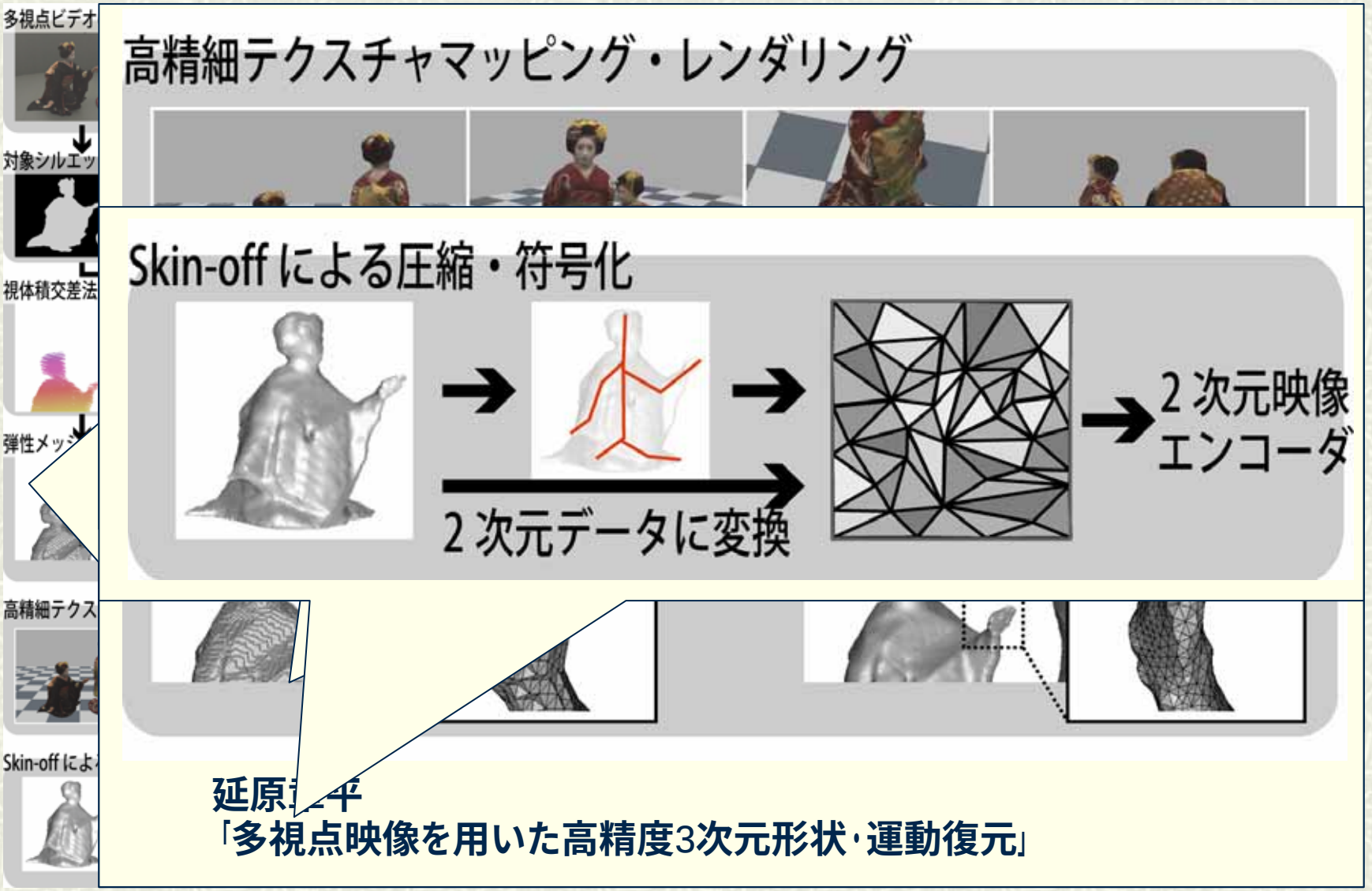
弾性メッシュ変形による
高精度3次元形状復元



高精細テクスチャマッピング

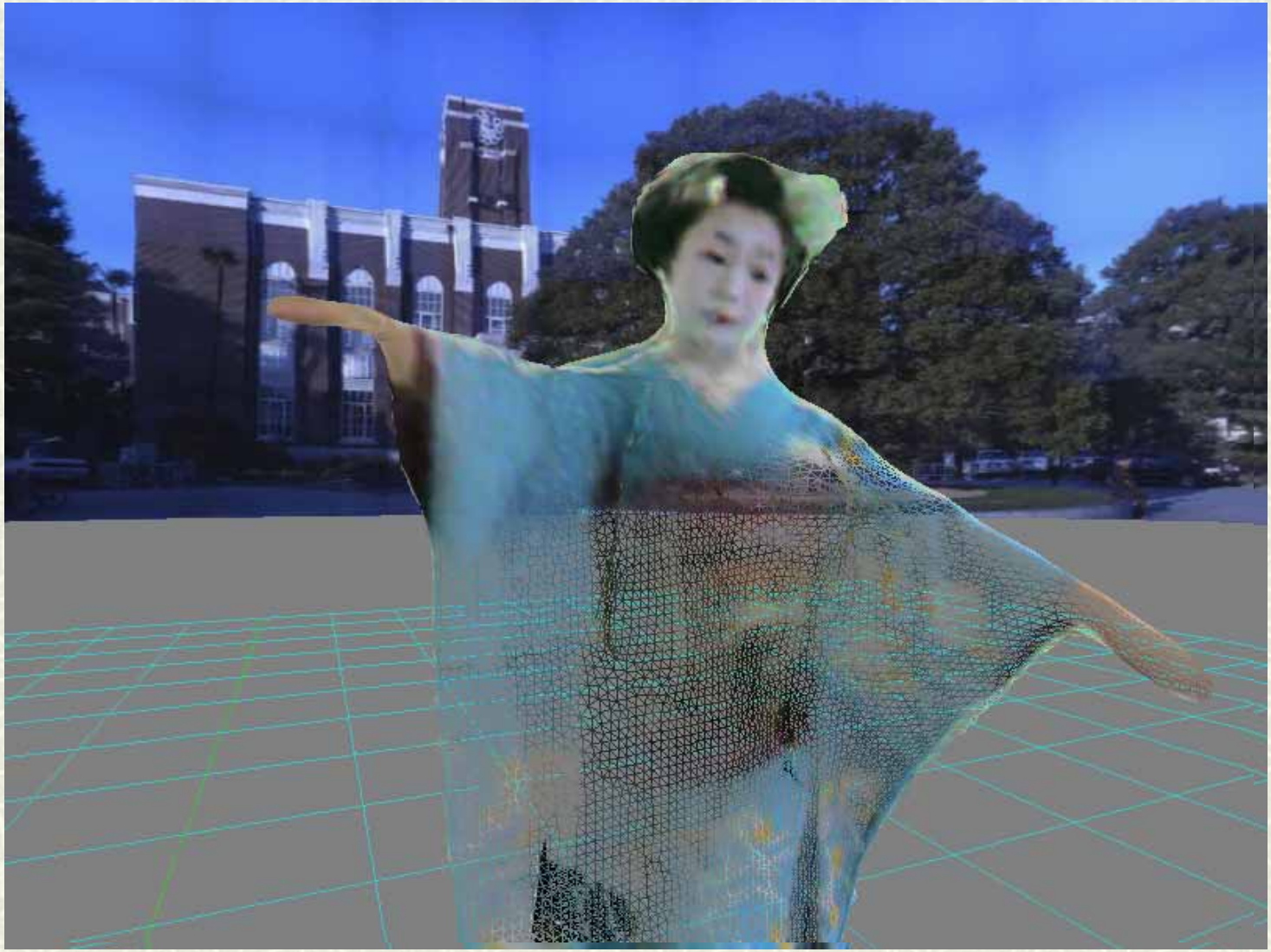
圧縮・符号化・伝送・表示

3次元ビデオ生成の流れ



延原[illegible]平
「多視点映像を用いた高精度3次元形状・運動復元」

「舞妓さん」の3次元ビデオの例



3次元ビデオの応用
（殺陣の振り付け）

53