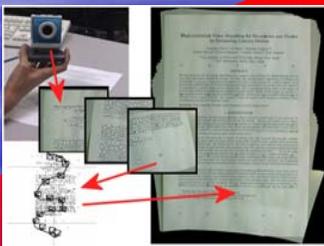


# 視覚情報メディア講座(横矢研究室)

教授 横矢直和, 准教授 山澤一誠, 助教 神原誠之, 佐藤智和

## コンピュータビジョン



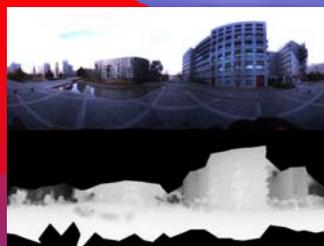
ビデオカメラによる超解像  
ビデオモザイクシステム



光学的整合性を考慮した  
拡張現実感システム



赤外線LEDと不可視マーカ  
によるカメラ位置・姿勢推定



全方位カメラによる  
屋外三次元情報の復元



パターン類似度に基づく  
画像修復



ネットワークを介した全方位カメラ  
による遠隔監視システム

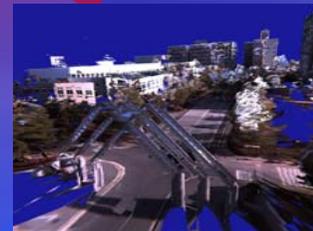
## 視覚・情報・メディア



ウェアラブル拡張現実感



ランドマークデータベース  
によるカメラ位置・姿勢推定



レーザレンジファインダによる  
屋外環境の三次元モデル化



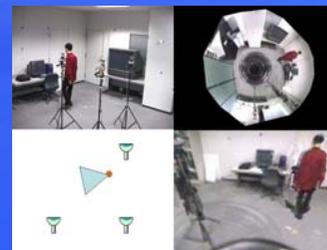
位置情報による  
撮影画像への索引付加



広域屋外環境の  
全方位ステレオ画像生成



画像のぼけ推定に基づく  
仮想物体の実時間レンダリング



全方位カメラによる  
新視点画像生成

## ヒューマンインタフェース

## コンピュータグラフィクス

# 視覚情報メディア講座 (VISION AND MEDIA COMPUTING LABORATORY)

教授:横矢 直和 准教授:山澤 一誠 助教:神原 誠之, 佐藤 智和

E-mail: (yokoya,yamazawa,kanbara,tomoka-s)@is.naist.jp

URL: http://yokoya.naist.jp/

## ■ 講座概要

本講座では、コンピュータやロボットが外界を「見る」ための技術とコンピュータ内部の多様な情報を人間に「魅せる」ための技術を中心に、**コンピュータビジョン**、**ヒューマンインタフェース**、**コンピュータグラフィクス**の分野とそれらの**複合領域**において視覚メディアの研究を行っています(図 1)。

## ■ 主な研究分野

### ・ コンピュータビジョン

画像を手掛かりにコンピュータやロボットに外部環境を理解させるための手法について研究しています。

- 動画像からの移動物体追跡、シーンの三次元復元(図 2)
- 実時間カメラパス推定に基づくビデオモザイク
- ランドマークデータベースによるカメラ位置・姿勢推定(図 3)
- レーザレンジファインダによる屋外環境の三次元モデリング(図 8)

### ・ ヒューマンインタフェース

コンピュータと人間の間の情報交換・インタラクションを容易にするための手法について研究しています。

- 全方位画像センサを用いたネットワークテレプレゼンス
- ウェアラブル拡張現実感システム(図 4)
- 装着した三次元センサによるユーザの自己位置計測

### ・ コンピュータグラフィクス

コンピュータを用いて現実に限りなく近い映像を再現するための手法について研究しています。

- 多重スケール環境マップを用いたレンダリング
- 画像のぼけ推定による仮想物体の実時間レンダリング(図 5)

### ・ 複合領域

上記の複数の分野の複合領域について研究しています。

- パターン類似度に基づく画像欠損領域の修復(図 6)
- 広域屋外環境の全方位立体映像の生成(図 7)
- 全方位カメラによる任意視点画像生成

## ■ 研究設備

- レーザレンジファインダ(図 8)
- 全方位センサ搭載インターネットカー(図 8)
- 画像処理サーバ(SGI Prism, Itanium2×16CPU) (図 8)
- 没入型ディスプレイ(傾斜型,球面型) (図 7, 図 8)
- 屋外センシング用インターネット飛行船(図 8)

## ■ 共同研究・社会活動など (平成 19 年度)

- 文部科学省科研費(基盤 A, 若手 B×2)
- 総務省(戦略的情報通信研究開発推進制度(SCOPE)×2)
- 科学技術振興機構(戦略的基礎研究推進事業(CREST), 地域イノベーション創出総合支援事業)
- 共同研究(トプコン, 豊田中研, オムロン)
- NEC 連携ラボ

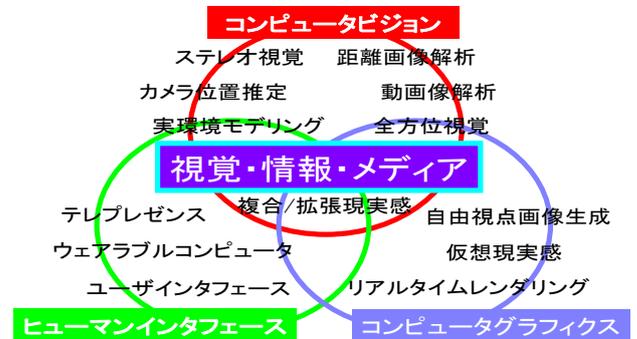


図 1: 視覚情報メディア講座の研究分野



図 2: 動画像からの三次元復元



図 3: 自然特徴による CG 位置合わせ



図 4: ウェアラブル拡張現実



図 5: 仮想物体のぼけ再現

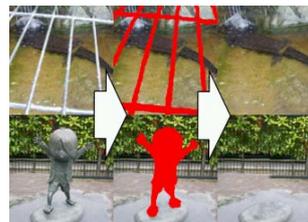


図 6: 画像欠損領域の修復



図 7: 全方位立体画像生成



画像処理サーバ

歩行装置付ディスプレイ

レーザレンジファインダ



屋外飛行船

全方位マルチカメラ

全方位センサ搭載車



移動ベース

54 眼カメラ

裸眼立体視モニター

図 8: 研究設備の一部

# 奥行き画像を利用した時空間超解像画像の生成

奈良先端科学技術大学院大学 視覚情報メディア講座  
粟津 優作, 河合 紀彦, 佐藤 智和, 横矢 直和

## 研究背景

一般のカメラでは高時空間解像度な映像の取得は困難  
(撮像素子の数が有限, 読み出し速度の制限等)

➡ 動画の時空間超解像化

### 従来研究の問題点

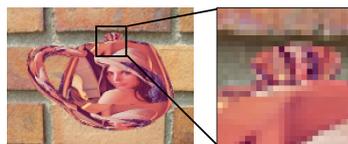
画像間の対応関係を用いて高解像度化  
高精度に対応関係を求めるために, 制約を要する

- 物体の形状を平面仮定
- カメラの動きを水平・垂直に限定

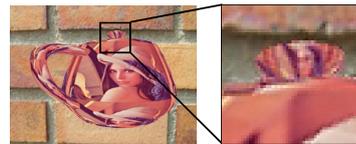
➡ 複雑な形状の物体等が対象の場合,  
良好な結果を得ることが困難

### 超解像処理

複数の画像などを用いて, より高解像度な画像を生成



入力画像



超解像画像  
(提案手法で生成)

## 目的とアプローチ

### 本研究の目的

対象物体の形状が複雑なシーンに対して撮影された動画画像に対応した時空間超解像画像の生成

### アプローチ

奥行き値を用いた1次元探索によって, 複数の観測画像間で各画素を高精度に対応付ける

画素独立に対応付けが可能

➡ 撮影対象に関する強い制約が不要

隣接フレーム間で対応点が存在しない場合にも,  
他のフレームを用いることが可能

➡ オクルージョンの影響を緩和

## 提案手法

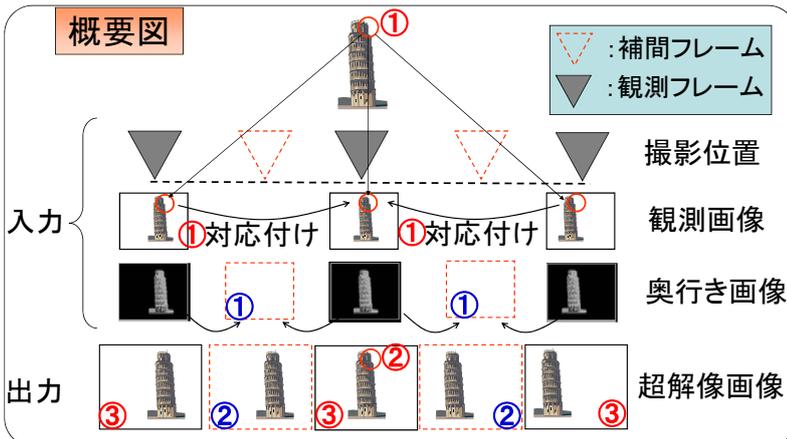
### 空間解像度の向上

- ①: 奥行き画像を用いてシーンの同一箇所に対応付け
- ②: 対応付いた画素をもとに超解像画像の画素を推定  
(推定に画素の整合性を考慮)
- ③: 全ての超解像画像の画素に同様の処理を施す

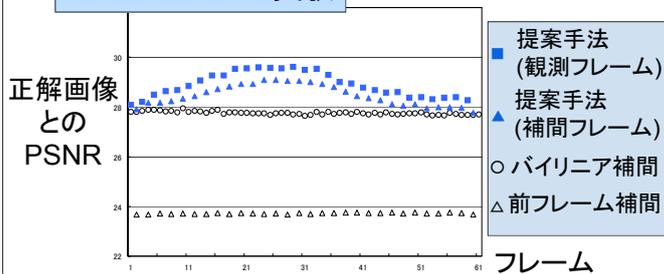
### 時間解像度の向上

- ①: 近傍フレームの奥行き画像をもとに補間フレームの奥行き画像を推定
- ②: 奥行き画像を用いて空間解像度の向上処理を適用

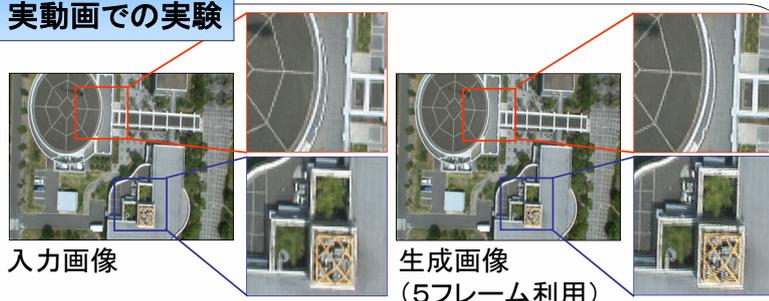
### 概要図



## シミュレーション実験



## 実動画での実験

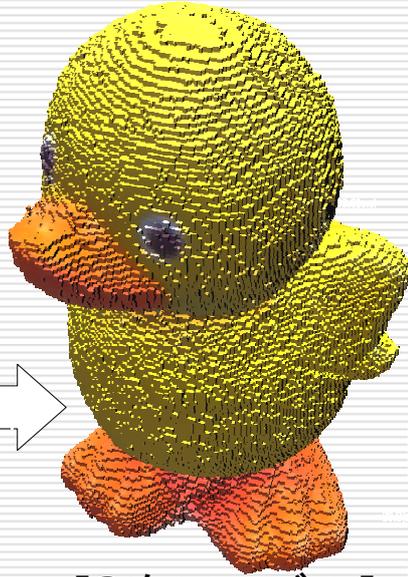


# ハンドヘルドビデオカメラを用いた 撮影位置指示インターフェースを有する インタラクティブ3次元モデリングシステム

視覚情報メディア講座 不殿健治, 佐藤智和, 横矢直和



[実物体]



[3次元モデル]

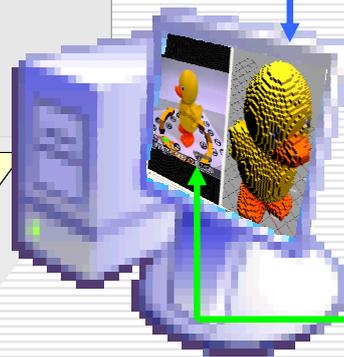
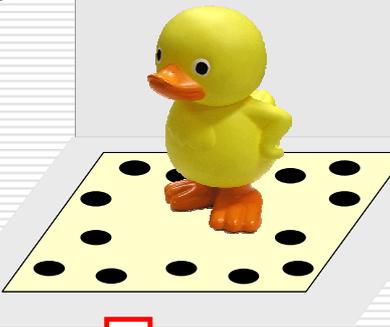
- ・個人のHP素材への利用
- ・オークションの商品説明への利用
- ・ゲームへの利用



手軽に3次元  
モデルを作りたい

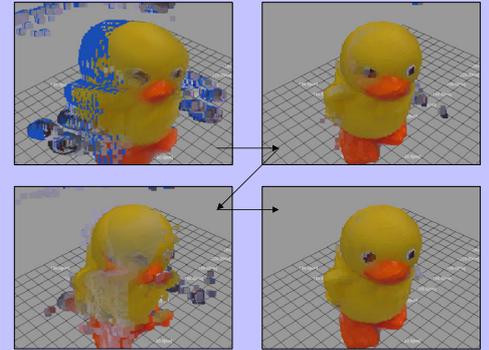
## インタラクティブ3次元モデリングシステム

撮影指示に従いながら撮影



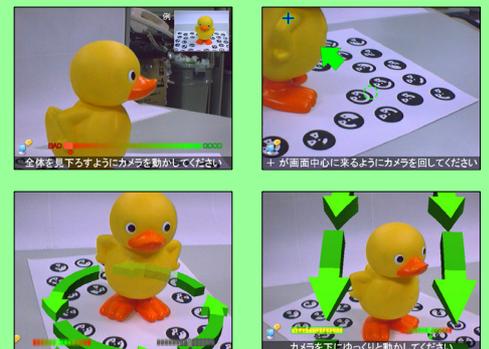
リアルタイムプレビュー

モデリング状況を随時確認



撮影指示

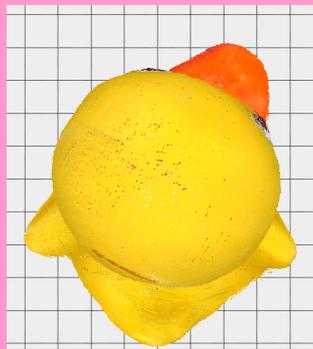
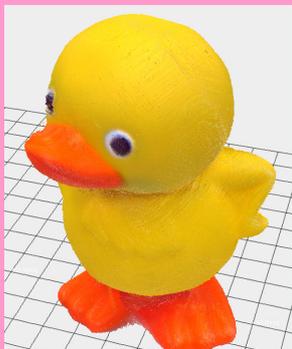
- ・基本的な撮影方法の教示
- ・撮りこぼしの無い効率的な撮影方法の教示



撮影終了

詳細モデリング(オフライン処理)

高精度なモデル生成



# ビデオカメラによる超解像 ビデオモザイクシステム

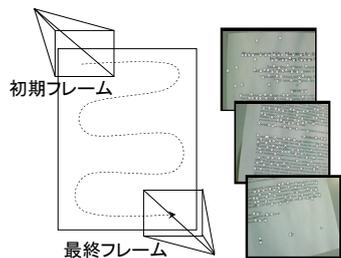
佐藤 智和, 池谷 彰彦, 池田 聖, 神原 誠之, 中島 昇, 横矢 直和  
視覚情報メディア講座 NEC産学連携ラボ

近年のインターネット・カメラ付き携帯電話やPDAの普及に伴い、**紙面に描かれた文章や写真を手軽に電子化**し、保存・伝送することに対する需要が高まっている。本研究では、ハンドヘルドカメラを用いて紙面をなぞるように撮影し、動画像解析を行うことで、カメラの位置・姿勢をリアルタイムで推定し、一枚の高解像画像を合成する。また、超解像処理を行うことで、出力画像の高解像化を図る。これにより、**個々のカメラの画質を大幅に改善する、高解像・広視野角の画像を生成できる。**

## 提案手法の処理の流れ:

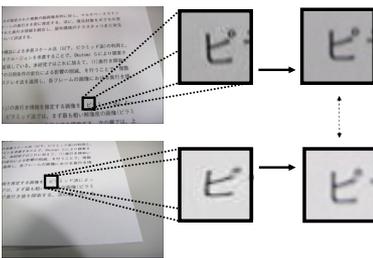
### 1. カメラ位置・姿勢のリアルタイム推定

特徴点の画像上での動きを解析することで、カメラの動きパラメータと特徴点のモザイク平面上の位置を推定する。



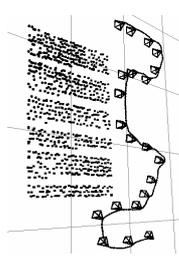
### 2. 再出現した特徴点の検出・対応付け

カメラの動きによってフレームアウトした後に再度撮影された特徴を、カメラパラメータを利用して検出し、対応付ける。



### 3. カメラパラメータの全体最適化

カメラパラメータと特徴点のモザイク平面上での位置を、動画像全体で最適化する。



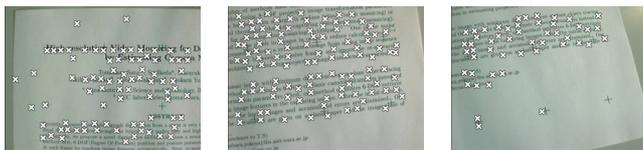
### 4. 超解像モザイク画像生成

最適化されたカメラパラメータを用いて超解像処理を行うことで、高解像モザイク画像を生成する。

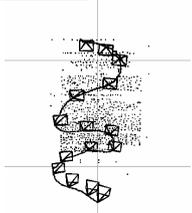


## 実験1 (文章)

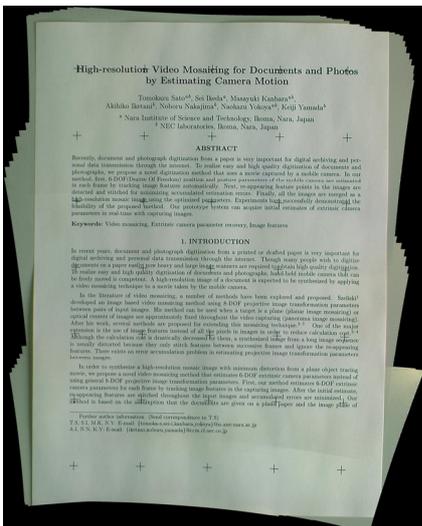
### 1. 入力画像・特徴点追跡結果の一部



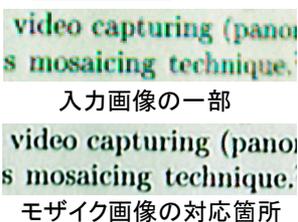
### 2. カメラパラメータの推定結果



### 3. モザイク画像の生成結果



### 4. 入力画像、モザイク画像の拡大図

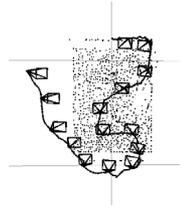


## 実験2 (写真)

### 1. 入力画像・特徴点追跡結果の一部



### 2. カメラパラメータの推定結果



### 4. 入力画像、モザイク画像の拡大図



### 3. モザイク画像の生成結果



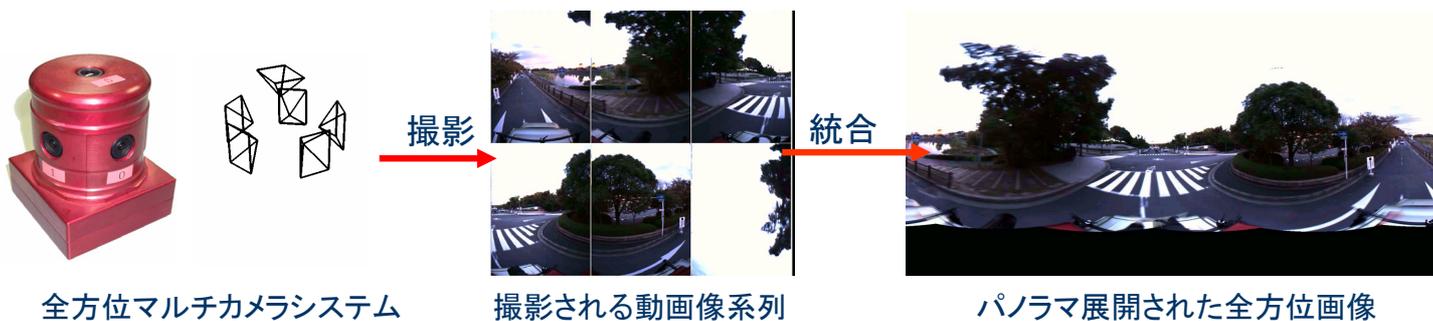
# 全方位カメラによる屋外三次元情報の復元

視覚情報メディア講座 佐藤 智和, 横矢 直和

<http://yokoya.naist.jp/~tomoka-s/research.html>

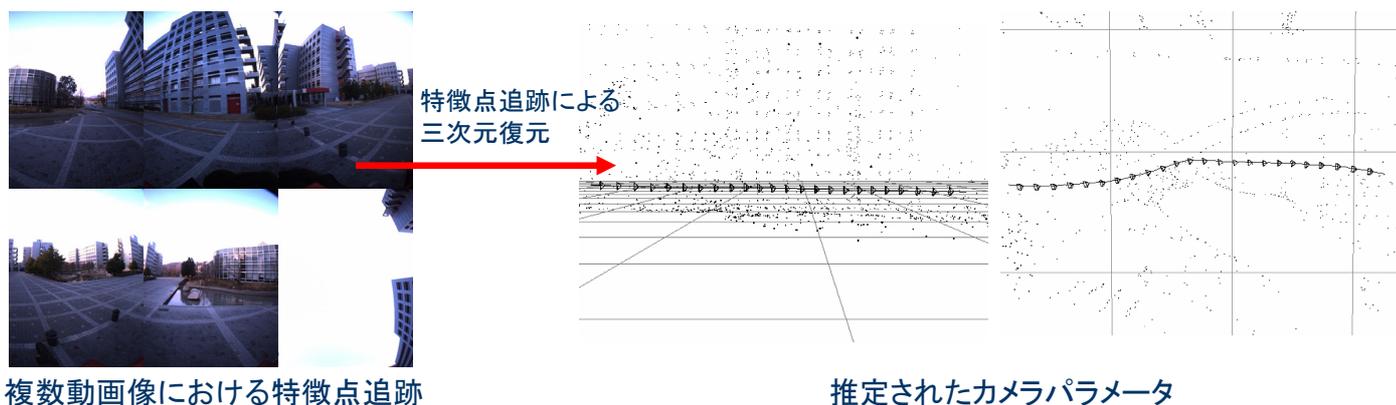
## 広域屋外環境の三次元モデル化

動画像からの屋外環境の三次元復元に全方位カメラを用いることで、三次元復元精度の向上と撮影コストの低減を図る。

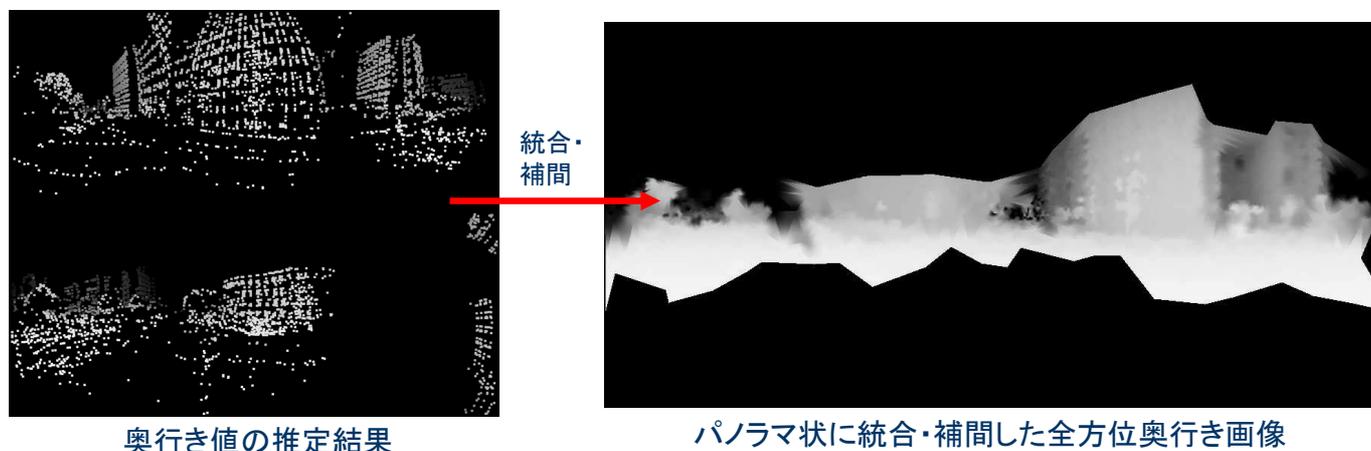


## 広域屋外環境の三次元情報の復元

1. 各カメラに写る特徴点を統合的に扱うことで高精度にカメラ位置・姿勢を推定



2. 特徴点の数え上げによる奥行き推定により、効率的に全方位の奥行きを復元



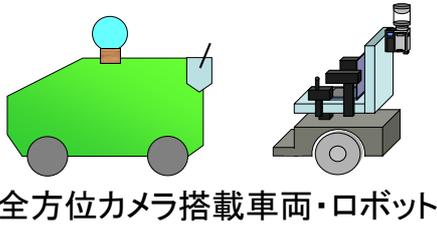
## 今後の展開

- ・ 全方位奥行き画像の統合による屋外三次元モデルの自動構築
- ・ 奥行き動画像を用いた任意視点画像生成手法の開発

# 複数の全方位動画像からの自由視点画像生成

視覚情報メディア講座 石川 智也, 山澤 一誠, 横矢 直和

動的実環境中を自由な視線方向・視点位置でテレプレゼンス



ロボットや車両に全方位カメラを搭載して遠隔操縦する方法

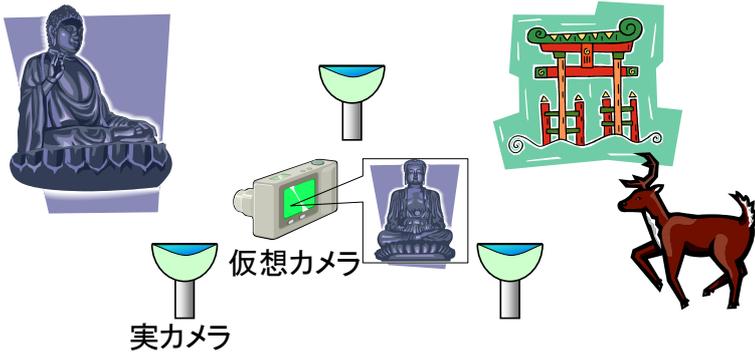
- ✖️ 視点移動に機械的な遅延が生じる
- ✖️ 一人の視点位置にしか対応できない

自由視点画像生成により仮想的に視点移動を行う方法

- 視点移動に機械的な遅延が生じない
- 複数人の視点位置に対応できる

アプローチ:

- 多地点全方位カメラによる実環境の撮影
- 静的領域と動的領域を分離して処理



多地点全方位画像を取得



前提条件

- カメラの位置・姿勢は既知
- カメラは固定して設置
- 各カメラは同期して動作

静的領域と動的領域を分離

静的領域

動的領域

画像間の対応付け

動的領域のみの画像を生成

背景画像を更新する背景差分処理

- フリッカーやカメラに依存するノイズを考慮
- 長時間動かない物体は静的領域と判断

仮想カメラの位置を設定

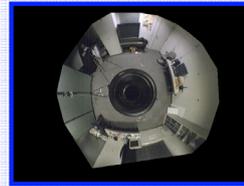
静的領域

動的領域

Morphingによる自由視点画像生成

Image-based VisualHullによる自由視点画像生成

静的・動的領域に対して自由視点画像生成

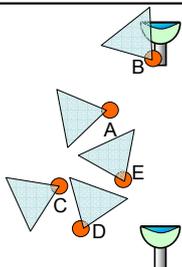


静的領域と動的領域の画像を合成



← 合成後の自由視点画像

画像提示



# 特徴点追跡とGPS測位に基づくカメラの位置・姿勢推定

横地 裕次, 池田 聖\*, 佐藤 智和\*, 横矢 直和\*  
 e-mail : \* {sei-i, tomoka-s, yokoya}@is.naist.jp

## 目的：カメラの位置・姿勢推定

- 広範囲に移動して取得した動画画像を使用
- 蓄積誤差の発生しない
- オフラインアプリケーション用
- カメラとGPS間のキャリブレーションが簡単
- 計算コストの小さい



画像とGPS測位データの取得

## 既知情報

- 画像とGPS測位データ
- カメラとGPS受信機間の距離
- カメラの内部パラメータ(レンズ歪, 焦点距離など)
- カメラとGPSの同期

## 未知情報

- カメラの位置・姿勢(外部パラメータ)
- 自然特徴点の三次元位置
- カメラからみたGPS受信機の方位

## GPS測位に基づくカメラの位置・姿勢推定アルゴリズム

### 特徴点追跡アルゴリズム

- (1) 特徴点の検出  
画像のスケール変化や回転に頑健なHarrisオペレータにより特徴点の候補位置を算出する
- (2) 特徴点の対応付け  
テンプレートマッチングにより第(i-1)フレームと第iフレーム上の特徴点に対応付ける
- (3) 暫定外部パラメータ推定  
RANSACアルゴリズムにより整合性の取れない特徴点を排除して暫定外部パラメータと特徴点の三次元位置を推定する
- (4) 特徴点の再対応付け  
暫定外部パラメータを用いて推定された特徴点の三次元位置を画像上に投影し、この位置の周辺のみ探索範囲を限定して、特徴点に対応付ける

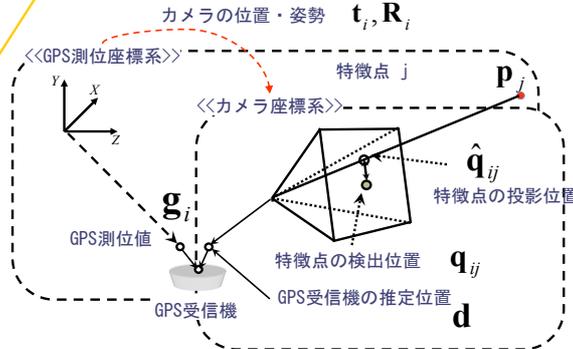
### 初期外部パラメータ

1フレーム内の特徴点の画像上の位置、三次元位置の対応から、そのフレームのカメラの位置・姿勢を推定する。カメラの位置・姿勢パラメータは、一般的に用いられる誤差関数である再投影誤差の自乗和を最小化することで算出する。

### 再投影誤差

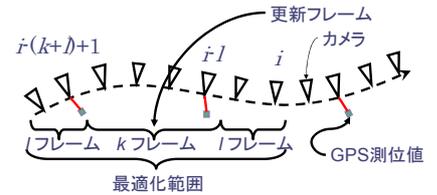
$$\Phi_{ij} = |q - \hat{q}|$$

### カメラとGPS受信機の幾何学的関係



### 狭区間最適化の範囲

更新フレームの最適化に十分な数の特徴点とGPS測位値を利用するために、更新フレームの両側に使い捨てる区間を設ける。



### 狭区間・広区間最適化の誤差関数

GPS受信機とカメラに関する誤差の重み  
第iフレームの信頼度

$$E = \frac{\omega}{|F|} \sum_{i \in F} \Psi_i^2 + \frac{1}{\sum_i |S_i|} \sum_i \mu_i \sum_{j \in S_i} w_j \Phi_{ij}^2$$

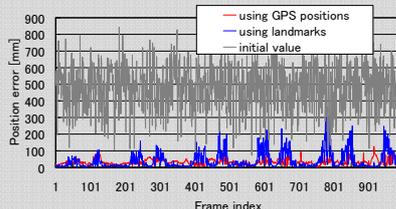
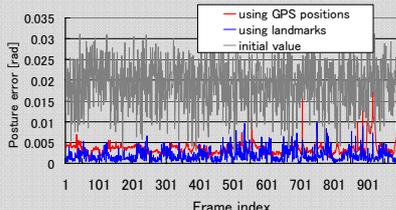
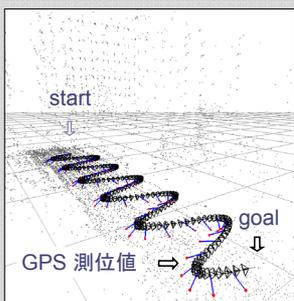
第iフレーム内の特徴点の集合  
特徴点jの信頼度  
GPS測位できたフレームの集合

### GPS受信機とカメラに関する誤差

$$\Psi_i = |R_i g_i + t_i - d_i|$$

## シミュレーションによる評価

### 処理(D)広区間最適化の精度評価



### 入力データ

- 画像 (720x480, 15fps) 990 フレーム
- GPS 測位値の精度 ( $\sigma=30\text{mm}$ )
- 特徴点検出誤差 ( $\sigma=0.6\text{mm}+\text{quantization error}$ )
- カメラとGPS受信機間の距離: 1039mm

定数  $k=5, l=22, \omega=10^{-9}$

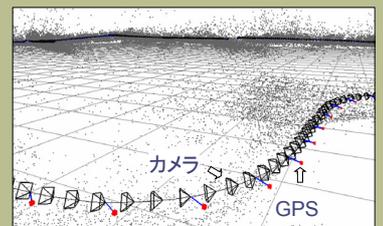
## 実環境データを用いた実験

誤差モデルや特徴点追跡精度を含む提案手法全体の評価

### 入力データ

- 画像 (720x480, 15fps), 3600 フレーム
- GPS 受信機の精度 ( $\sigma=30\text{mm}$ )
- カメラとGPS受信機間の測定距離: 1020mm

定数  $k=5, l=22, \omega=10^{-9}$



[http://yokoya-lab.naist.jp/pub/movie/yokochi/match\\_move.mpg](http://yokoya-lab.naist.jp/pub/movie/yokochi/match_move.mpg)

# ウェアラブル拡張現実感のための 不可視マーカと赤外線カメラを用いた 位置・姿勢推定システム

奈良先端科学技術大学院大学 視覚情報メディア講座 中里 祐介  
URL: <http://yokoya.naist.jp/~yuusu-n/>      mailto:yuusu-n@is.naist.jp

ヒューマンナビゲーション等のウェアラブル型拡張現実感システムには

**ユーザの正確な位置・姿勢情報が必要**

- ◆屋外: 一般的にGPS + ジャイロセンサを使用
- ◆屋内: GPSが使用できないため、広域で容易に使える位置計測手法が必要



ウェアラブル拡張現実感システム

## 目的とアプローチ

屋内の位置検出において

インフラを簡単に設置したい

**画像マーカ**



予め位置を計測した画像マーカを認識することにより、ユーザの位置・姿勢を推定

ただし、画像マーカは景観を損ねてしまう

景観を損ねたくない

再帰性反射材からなる**不可視マーカ**



光源方向に光を反射するため、光源の近くにあるカメラのみがマーカを撮影可能



フラッシュ無しで撮影



フラッシュ有りで撮影

## システム構成

### 環境構築

管理者

不可視マーカの準備と貼り付け

不可視マーカのキャリブレーション



不可視マーカ壁紙

- ・正方形マーカを密に配置
- ・パターン: 正方形枠内を  $n \times n$  の格子状に分け、点を配置  
内部パターンの四隅の点は1つのみ  
CビットのCRC (Cyclic Redundancy Check) コードを付加  
パターン数  $2^{n^2+c}$  通り



不可視マーカの撮影

キャリブレーションツールによるマーカ配置情報作成



入力  
・基準マーカの位置・姿勢  
・マーカ撮影画像  
・カメラ内部パラメータ  
・不可視マーカ設計情報

出力  
・マーカ配置情報:  
各マーカの3次元位置・姿勢

マーカ配置情報 (各マーカの3次元位置・姿勢)

### ユーザの位置・姿勢推定

ユーザ

ユーザ位置・姿勢推定システムの実行

ユーザ位置・姿勢

(UDP送信 or 共有メモリに出力)

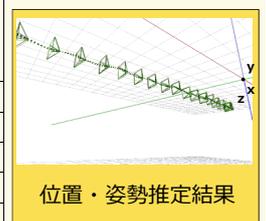
不可視マーカ



撮影画像

赤外線LED付き赤外線カメラ

	赤外線カメラA	赤外線カメラB
特長	高精度な位置推定	小型・軽量
本体重量	115 g	20 g
解像度	1024 × 768 画素	640 × 480 画素
水平画角	110°	70°
その他	前方撮影用カメラ付き	
位置・姿勢計測精度	9mm, 0.4°	15mm, 0.5°



位置・姿勢推定結果

# ウェアラブル拡張現実感システムのための位置依存情報オーサリング

視覚情報メディア講座（横矢研究室）  
Web: <http://yokoya.naist.jp>

## 研究背景

### ウェアラブル拡張現実感システム

- ユーザの位置・姿勢に応じて位置依存情報をユーザの眼前の風景に重畳表示
- ヒューマンナビゲーション、景観シミュレーションなどに応用可能



ウェアラブルPCのユーザ

ユーザに提供する位置依存情報の管理・提供（オーサリング）が必要

## 位置依存情報のオーサリングのアプローチ

### (1) 3Dモデルの設定

地図情報と拡張現実感システムを利用した3Dモデルの位置・姿勢の設定

### (2) 注釈情報の設定

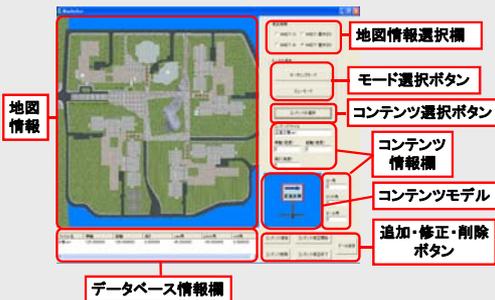
地図情報と拡張現実感システムを利用した注釈対象の形状設定

### (3) 位置依存情報の提供

ネットワーク共有データベースを利用した情報の提供

## (1) 3Dモデルの位置・姿勢設定

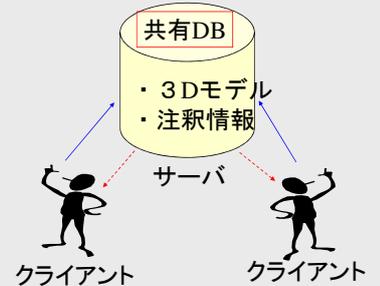
第1フェーズ：通常のPCによる地図情報を利用した効率的な情報設定



## (2) 注釈対象の形状設定



## (3) 位置依存情報の提供



ユーザID・自己位置  
ユーザ位置付近の3Dモデル及び注釈情報

**サーバ:** 環境中の3Dモデル、注釈情報及び全ユーザの現在位置を管理

**クライアント:** 自己位置付近の3Dモデル及び注釈情報を獲得しユーザに提示

位置依存情報の効率的な管理及びクライアントへの情報配信が可能

第2フェーズ：拡張現実感システムを利用した情報の詳細設定



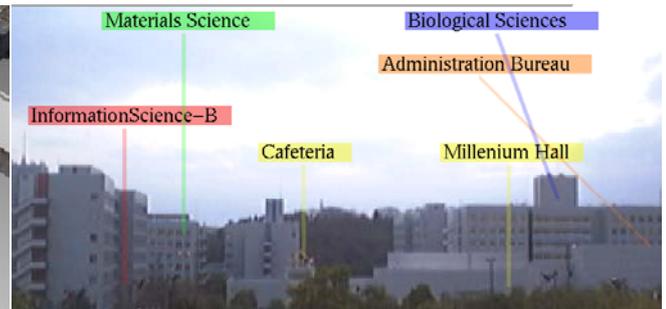
## 生成画像例



3Dモデルの提示画像



注釈対象の形状設定結果



注釈提示画像

# 画像特徴点によるランドマークデータベースに基づくカメラ位置・姿勢推定

○大江 統子, 佐藤 智和, 横矢 直和  
奈良先端科学技術大学院大学 視覚情報メディア講座

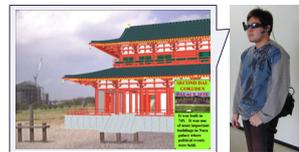
## 研究目的とアプローチ

拡張現実感やロボットナビゲーションではカメラの位置・姿勢が必要とされる  
本研究では、

**複雑で広範囲な環境におけるカメラの絶対位置・姿勢を推定**



- 画像内の自然特徴点周辺の画像をランドマークとしてマーカの代わりに利用
- 全方位動画像を用いることで事前に多数のランドマークを作成



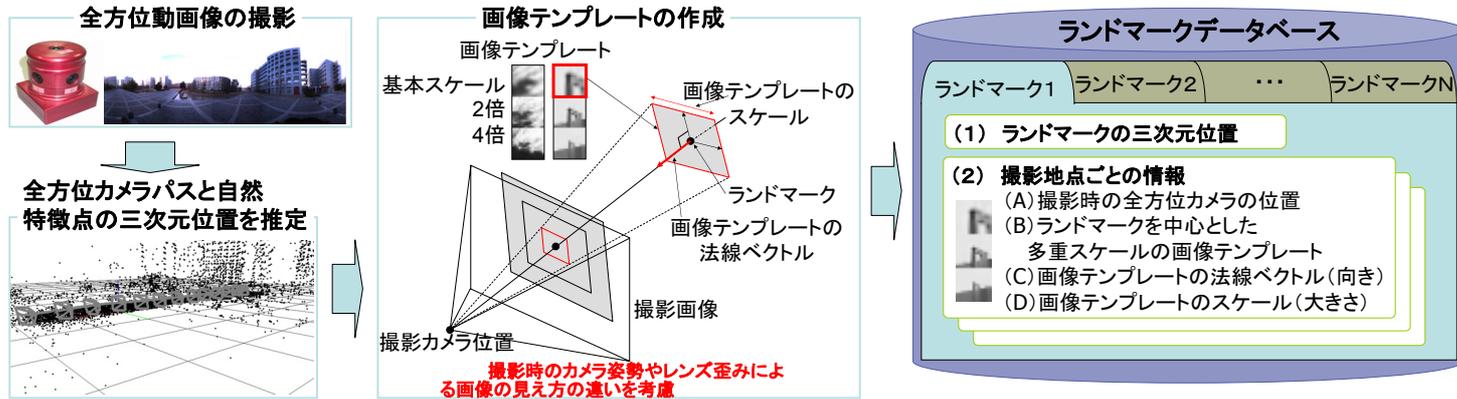
ウェアラブルARシステム[天目]

### <特徴>

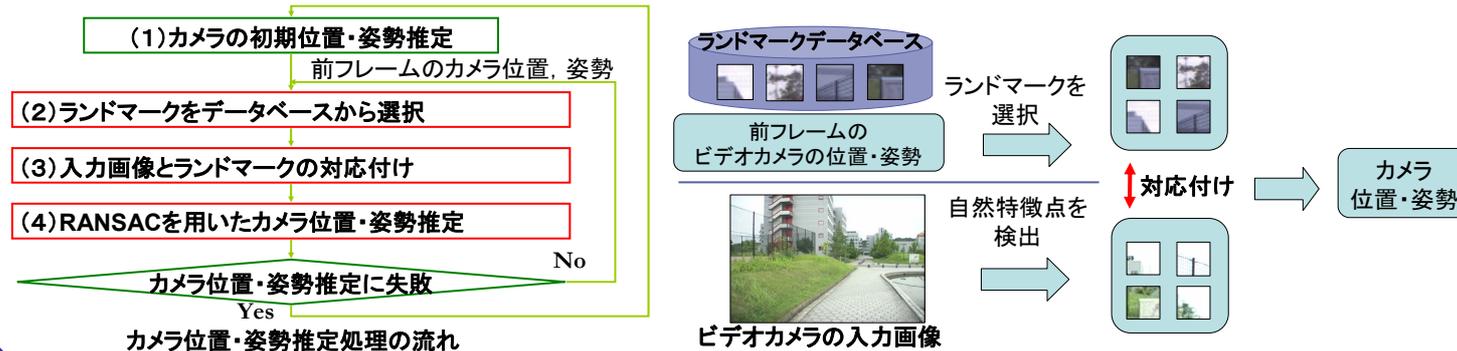
- カメラ以外のセンサや人工的なマーカを用いない
- 広範囲な屋内外環境で利用可能
- カメラ位置・姿勢推定時に累積誤差を含まない

## 提案手法の概要

### (1) 環境内の自然特徴点によるランドマークデータベースの構築(オフライン処理)



### (2) ランドマークデータベースを用いたビデオカメラ位置・姿勢の推定(逐次処理)



## 実験結果

- データベースから選択されたランドマーク
- ⊗ 入力画像と対応付けられたランドマーク
- RANSACによって誤対応を排除した後のランドマーク

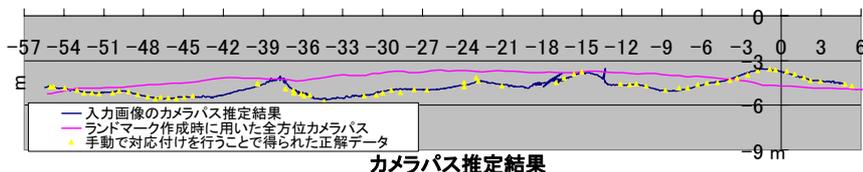


カメラ位置・姿勢推定に用いられたランドマーク



カメラ位置・姿勢推定結果を用いたmatch move

カメラ位置・姿勢推定に要した時間: 1フレームあたり2~3秒  
カメラ位置・姿勢推定に用いたランドマーク数: 1フレームあたり100個  
マシンスペック: CPU Pentium4 3.0GHz メモリ 1.5GB



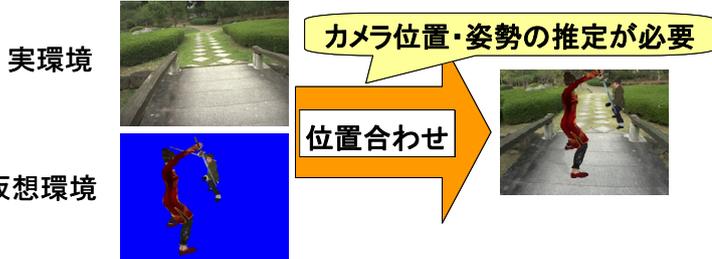
カメラパス推定結果

# 優先度情報を付加したランドマークデータベースに基づく実時間でのカメラ位置・姿勢推定

視覚情報メディア講座 武富 貴史, 佐藤 智和, 横矢 直和

## 拡張現実感とは

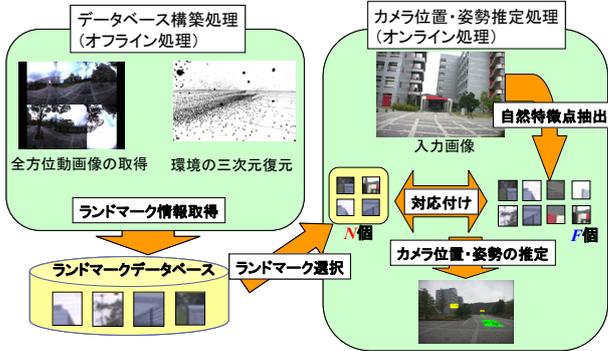
現実環境を撮影した画像に対してCGなどを位置合わせして描画することで情報を付加する技術



## カメラ位置・姿勢推定の要求事項

- 絶対位置・姿勢を取得できる
- 誤差が蓄積しない
- 実時間で処理できる

## ランドマークデータベースに基づくカメラ位置・姿勢推定



- データベースの構築が比較的容易
- 広域な環境で蓄積誤差を含まないカメラの絶対位置・姿勢の推定が可能

### 問題点

ランドマークと入力画像中の自然特徴点の照合に多くの計算コストを必要とし、実時間での処理が困難

## 研究目的

ランドマークデータベースを用いた実時間でのカメラ位置・姿勢推定の実現

## アプローチ

ランドマークと自然特徴点の照合回数を削減することで対応付け処理の計算コストを低減

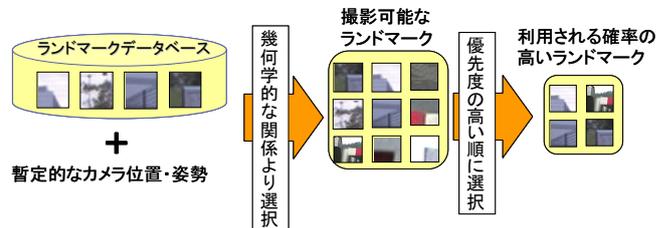
### 照合すべき自然特徴点数の削減

連続フレーム間でのランドマーク追跡処理によって、現フレームの暫定的なカメラパラメータを推定し対応点の探索範囲を小領域に限定

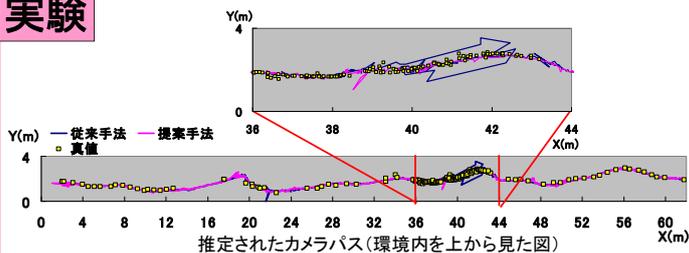


### 照合に用いるランドマーク数の削減

ランドマークへ優先度情報を付加することで、正しく対応づく可能性の高い小数のランドマークをデータベース中より選択



## 実験



位置・姿勢推定精度の比較

	従来手法	提案手法
平均位置誤差(mm)	338(標準偏差555)	233(標準偏差147)
カメラの光軸向きの平均誤差(度)	0.80(標準偏差0.62)	0.81(標準偏差0.60)

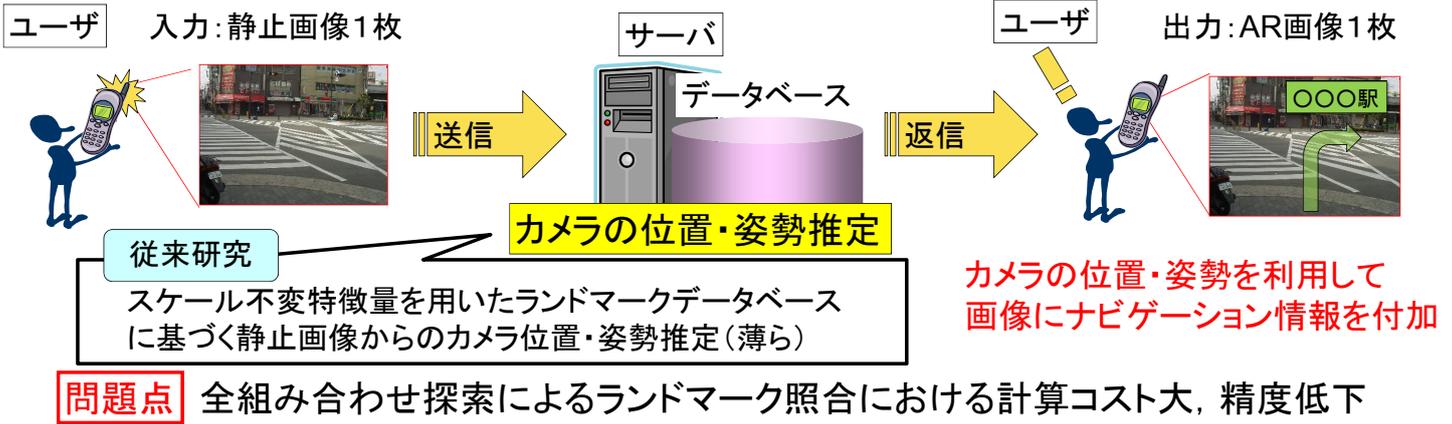
処理時間の比較

	処理時間 (ms)		処理時間の比
	従来手法 (ランドマーク数80個)	提案手法 (ランドマーク数60個)	
暫定的なカメラ位置・姿勢の推定	-	28	-
ランドマークの選択	12	1	0.08
ランドマークの対応付け	316	15	0.05
誤対応の排除とカメラ位置・姿勢推定	61	17	0.28
画像の取得など	4	5	1.25
全体	393(2.5fps)	66(15.1fps)	0.16

# 画像特徴点の特異度を考慮した ランドマーク照合による一枚の画像からの カメラ位置・姿勢推定の高速化とロバスト性の向上

奈良先端科学技術大学院大学 視覚情報メディア講座  
西海 嘉志, 佐藤 智和, 横矢 直和

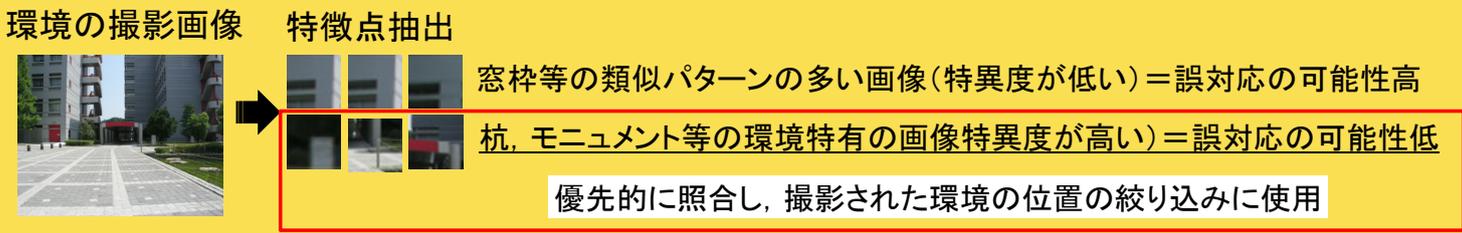
## カメラ付き携帯機器を利用したヒューマンナビゲーション



**研究目的** ランドマーク照合における計算コストの低減, 誤対応の低減

**アプローチ** 環境における特異さを用いたランドマーク照合に用いるランドマークの絞り込み

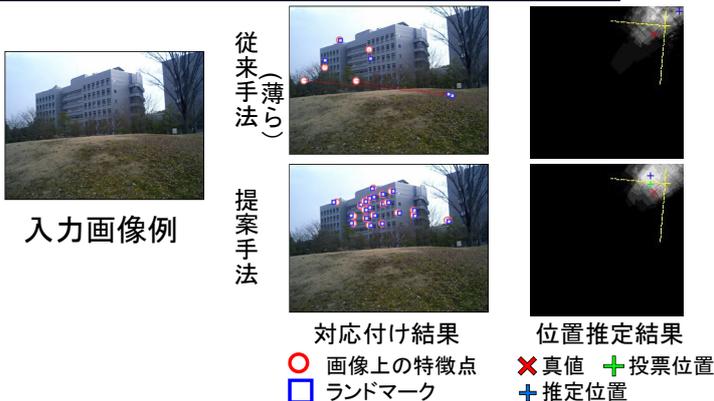
**特異度** 対応付けの信頼度の指標として設定



## 処理の流れ



## 実験: 実環境においてデータベースを作成



	従来手法(薄ら)	提案手法
推定成功率	41.2%	52.7%
平均位置誤差	6.8m	2.8m
平均姿勢誤差	3.9m	1.2m
平均再投影誤差	2.0画素	2.2画素

処理	従来手法(薄ら)	提案手法
特徴抽出	14.5	14.7
特異なランドマークの選択	-	3.8
ランドマークの選択	111.6	35.7
カメラ位置・姿勢推定	10.3	2.8
合計[sec]	136.4	57.0

# 全方位マルチカメラシステムを用いた蓄積再生型テレプレゼンス

視覚情報メディア講座 池田 聖, 佐藤 智和, 横矢 直和

遠隔地の映像をインタラクティブに提示することにより、あたかも遠隔地にいるかのような感覚を与える技術はテレプレゼンスと呼ばれる。複数のカメラを組み合わせた、全方位型のマルチカメラシステムを用いて高解像度な全方位動画像を取得し、没入型ディスプレイに提示することで臨場感の高いテレプレゼンスを実現する。全方位マルチカメラシステムからは、レンズ歪み、周辺減光、カメラ間の色調のずれを含んだ各カメラの画像が得られる。あらかじめ画像からカメラシステムの内部構造やレンズ・撮像素子の特性を推定し、自動的にそれらを補正しつなぎ合わせた画像を生成する。このようにして生成された画像と没入型ディスプレイを用いた2種類のテレプレゼンスシステムにより遠隔地の疑似体験ができる。

## ① 画像の取得:

### 全方位マルチカメラシステム Ladybug



カメラヘッド 蓄積ユニット



### 特徴

- 水平5個+上向き1個のビデオカメラを装備
- 一度に全天球の約75%の視野を撮影可能
- 高解像度 (768×1024×6)
- 動画像 (15 fps) として取得可能
- 約20分間蓄積可能

## ② 提示画像の生成:

### 光学的キャリブレーション

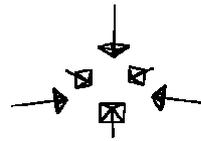


補正前 補正後 補正前 補正後  
① 周辺減光補正 ② 色調補正

### 幾何学的キャリブレーション



補正前 補正後  
③ 歪み補正



カメラの位置・姿勢



つなぎ合わされた全方位画像

④ カメラの位置・姿勢を用いた画像のつなぎ合わせ

## ③ 没入型ディスプレイによる提示:

### 動的な環境のテレプレゼンス

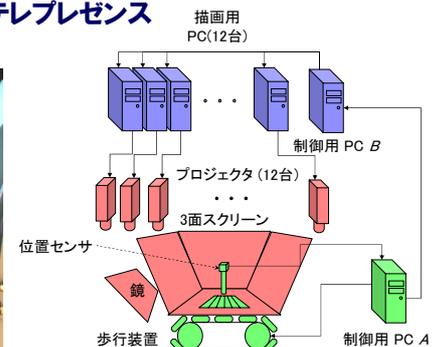


球面ディスプレイとゲームコントローラを用いたテレプレゼンスシステム。移動する自動車やジェットコースターに乗り、周囲の景色を見回す感覚を得ることができる。利用者はゲームコントローラを用いた見たい方向を自由に選択する。

### 静的な環境のテレプレゼンス



歩行装置とマルチプロジェクションディスプレイを用いたテレプレゼンスシステム。周囲を見回しながら静的な遠隔環境を自由に歩行移動する感覚を得ることができる。利用者の両足に取り付けられた、位置センサの情報を基に歩行装置のベルト及び映像が制御される。



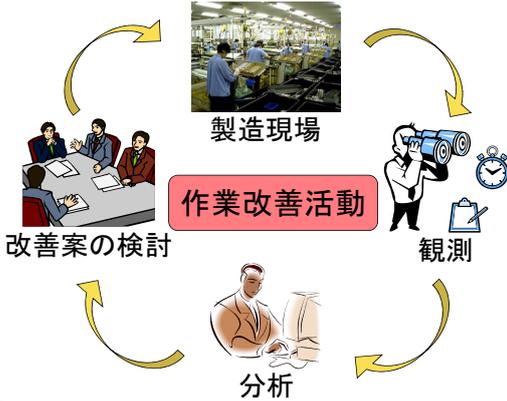
# 人物大装置の組み立て作業改善活動支援のための 距離画像センサを用いた俯瞰画像からの人物姿勢分類

高橋 典宏\*1, 山澤 一誠\*1, 生雲 公啓\*2, 野田 賢\*3, 横矢 直和\*1

奈良先端科学技術大学院大学 {\*1視覚情報メディア講座, \*3システム制御・管理講座} \*2オムロン株式会社

## 研究背景

製造工場では、生産性・製品品質向上のための作業改善活動(IE: Industrial Engineering)が常に行われている



### 従来のIE

- ・ストップウォッチを用いて作業時間を計測
- ・手書きによる作業内容, 作業時間の記録
- ・同じ作業の繰り返し観測が必要

観測・分析では多くの手間と時間がかかる

効率化

IE支援手法の研究・開発に期待

### 従来のIE支援手法



ジャンシ株式会社 "OTRS" など

○ 観測・分析の効率化と精度向上を実現

✕ リードタイムの長い作業への適用が困難

※リードタイム: 製造の開始から終了までに要する時間

## 研究目的

作業改善活動支援のための人物の姿勢分類

### 対象作業

セル生産方式による人物大装置の組み立て作業  
券売機, ATM, 基板検査装置など

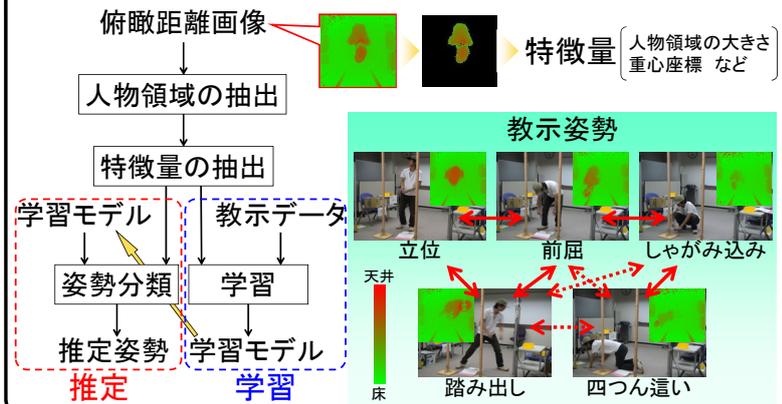


リードタイムが長い

- ✕ 観測漏れのないIEは困難
- 複雑な多数の工程を持つ
- ✕ 同工程, 同作業でも動作に再現性がない

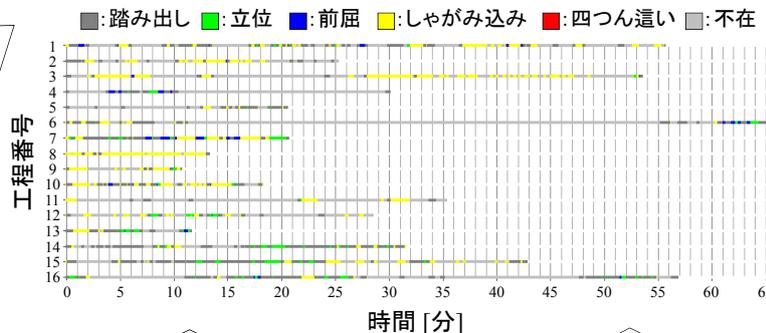
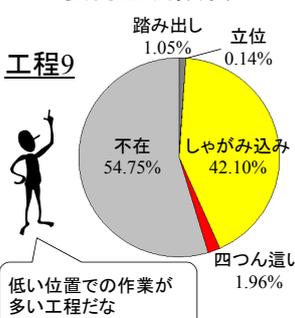
※セル生産方式: 1~数人の作業者がセル(作業場)内において, 一人が多工程を担当しながら部品の取り付けから検査までの全工程を行う生産方式

## 俯瞰距離画像からの人物姿勢分類

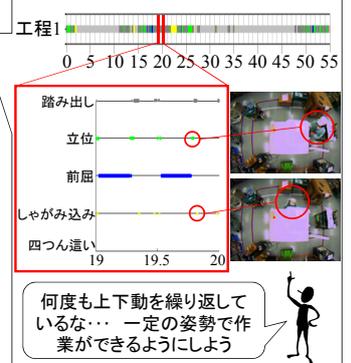


## 作業者の姿勢遷移に基づくIE支援

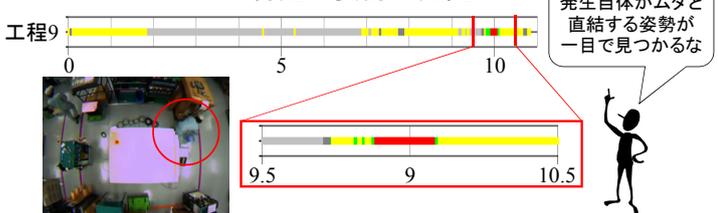
### 単位工程(時間)当たりの姿勢出現頻度



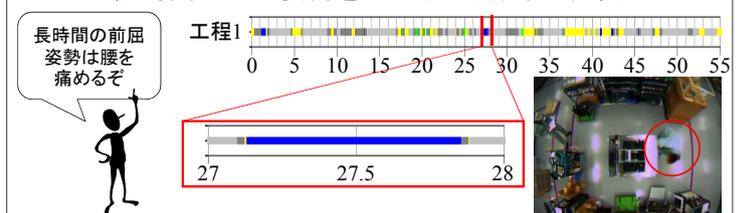
### 特定の姿勢遷移



### 特定の姿勢の発見



### 長時間同一の姿勢を取っている作業の発見



# 地理情報データベースとweb検索を用いた 撮影位置・姿勢情報に基づく 写真キャプションニングシステム

岩崎 季世子, 山澤 一誠, 横矢 直和

URL: <http://yokoya.naist.jp/~kiyoko-i/> E-mail: [kiyoko-i@is.naist.jp](mailto:kiyoko-i@is.naist.jp)

## 研究背景

デジタルカメラ等の手軽な  
撮影機器の普及  
大量の写真が  
整理されないまま蓄積



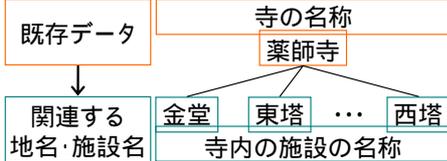
キャプションにより  
効率的な閲覧・検索を実現



## 撮影位置・姿勢情報に基づく半自動写真キャプションニングシステム

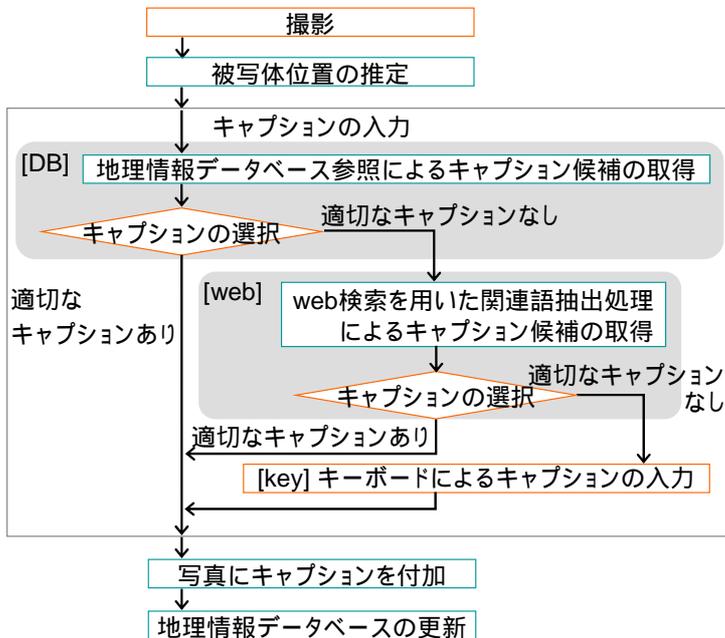
## アプローチ

- 撮影位置・姿勢情報に基づくキャプション候補の取得
  - 地図データベースの参照
  - web検索を用いた関連語抽出処理



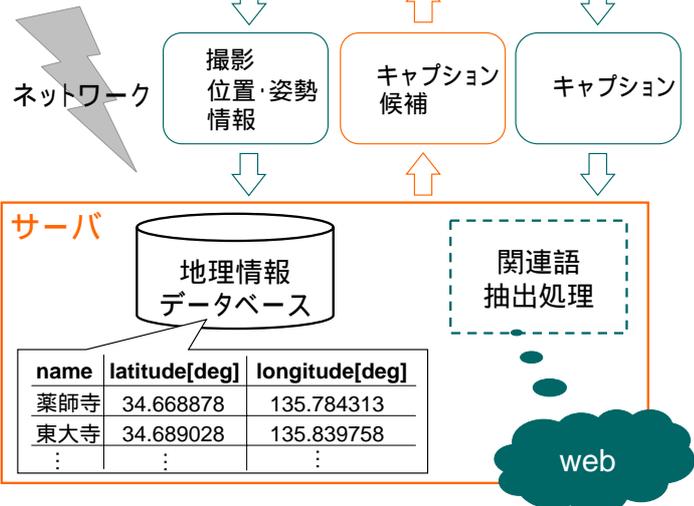
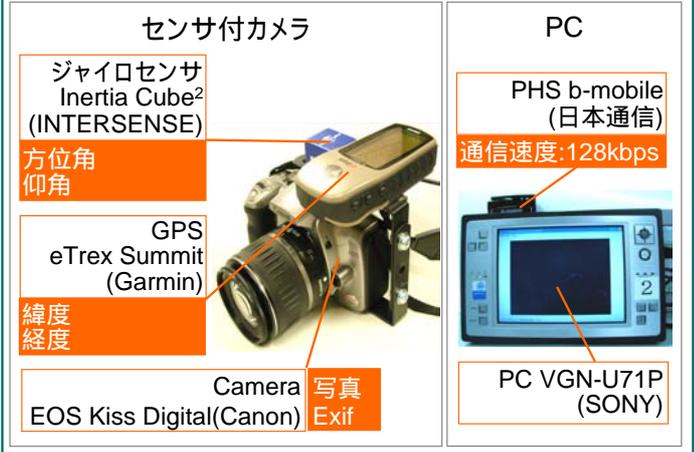
- ユーザによるキャプション選択をフィードバックとした  
地理情報データベースの更新

## 処理の流れ



## プロトタイプシステム

### クライアント



## 実験

ユーザ12名が各4枚の写真撮影・キャプションニング

- [撮影対象] 学内の施設
- 情報科学研究科: 2枚
- バイオサイエンス研究科: 1枚
- 他4施設: ユーザ任意の1枚
- [地理情報DB]
- 市販の地図ソフトの施設データを予め登録  
(撮影対象施設のうち1施設のみ登録済)

実験全体で  
48枚



ユーザによる入力の様子  
キャプション選択

### キャプションニング作業結果

キャプション 入力方法	キャプションを 付加した枚数	平均入力 時間[秒]	平均提示 順位
[DB]	43	10	1.8
[web]	3	22	66.3
[key]	2	113	-

ユーザのシステム利用による地理情報データの新規登録・更新 → キャプションニング作業の効率化(より適切な候補の提示)

# 拡張現実感のための画質に着目した 幾何学的・光学的整合性の向上に関する研究

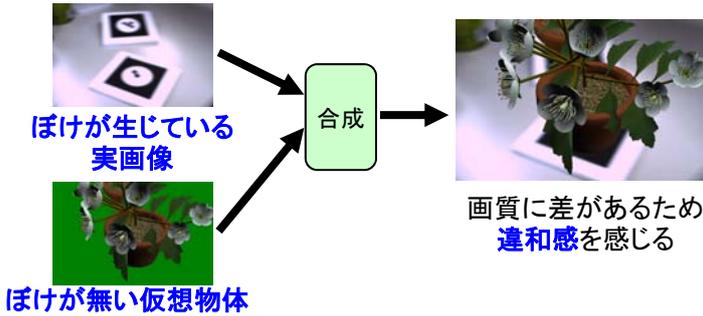
奈良先端科学技術大学院大学 視覚情報メディア講座  
 ○奥村 文洋 神原 誠之 横矢 直和  
 {bunyo-o, kanbara, yokoya}@is.naist.jp URL: http://yokoya.naist.jp/

## 拡張現実感における画質に関する整合性問題

拡張現実感における基本的な問題:

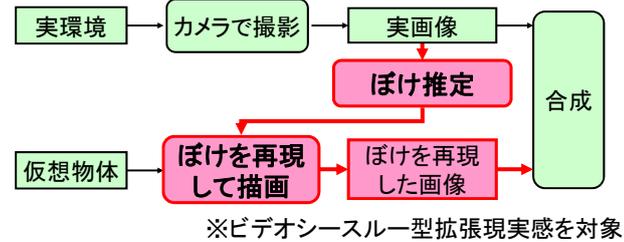
- 幾何学的整合性 (実環境と仮想環境の位置あわせに関する整合性)
- 光学的整合性 (仮想物体と実環境の陰影や画質に関する整合性)

従来、実画像と仮想物体の画質の差はあまり考慮されていない



## 実環境と仮想物体の間の画質の差を軽減するためのアプローチ

実画像からのぼけ推定と仮想物体に対するぼけ再現に基づいた実時間画像合成



### 【特徴】

- パラメータ表現されたPSFによる画像のぼけの記述
  - PSFパラメータをマーカのエッジから実時間推定
  - 仮想物体に対するぼけ再現をグラフィックスハードウェアを用いて高速化
- (PSF: Point Spread Function)

## 画質の一致を図った画像合成手法

### 提案手法で用いるぼけモデル

#### ● 焦点ずれによるぼけ ( $\hat{z}_f$ )

- カメラから物体までの奥行きに依存軸
- 合焦位置

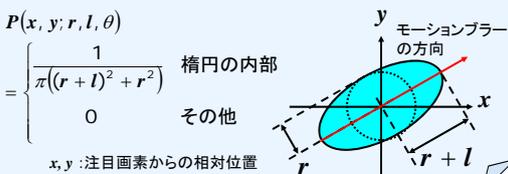
#### ● モーションブラー ( $\hat{l}, \hat{\theta}$ )

- 画像面上で均一な平行移動 (ユーザの頭部の回転によって生じると仮定)
- モーションブラーの大きさ, 方向

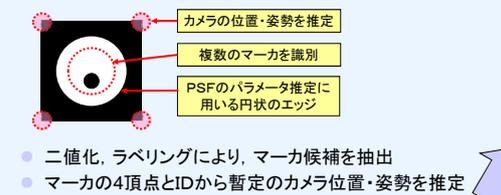
### 提案手法で用いるPSFの形状

$$P(x, y; r, l, \theta) = \begin{cases} \frac{1}{\pi(r+l)^2 + r^2} & \text{楕円の内部} \\ 0 & \text{その他} \end{cases}$$

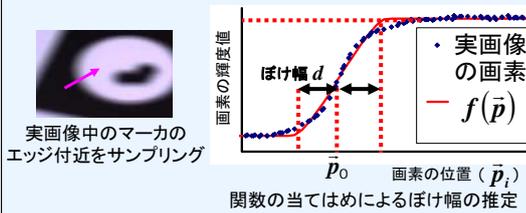
$x, y$ : 注目画素からの相対位置



### 処理A-提案手法で用いるマーカとマーカの検出



### 処理B1-マーカのエッジ付近の画素に対する関数当てはめによるぼけ幅の推定



### 処理B2-様々な方向のエッジ毎のぼけ幅からのPSFパラメータ ( $r, l, \theta$ ) の推定



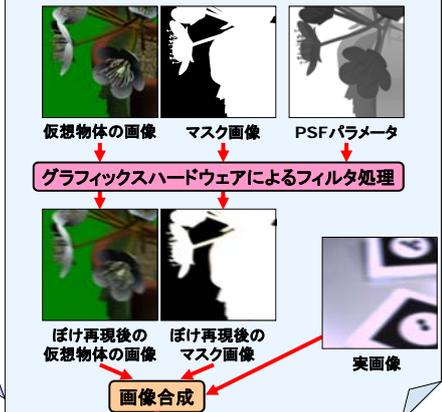
### 処理C-ぼけを考慮したカメラ位置・姿勢の推定

1. マーカの外周エッジの位置の推定
2. 外周エッジの位置に対する直線の当てはめ
3. 直線の交点からマーカの頂点を推定
4. マーカ座標系におけるカメラ位置・姿勢の推定

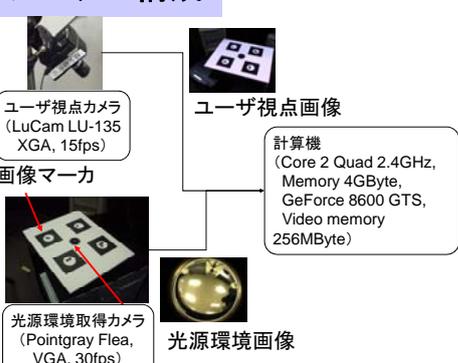
### 処理D-カメラの合焦位置とモーションブラーの推定

奥行き異なる複数のマーカのPSFパラメータから推定  
 合焦位置: 評価関数 (PSFのパラメータ  $r$  とぼけと奥行きの関係式の差の2乗和) が最小となる合焦位置  
 モーションブラー: 各マーカのPSFパラメータ  $l, \theta$  の平均値

### 処理E-ぼけを考慮した仮想物体の描画と実画像への合成

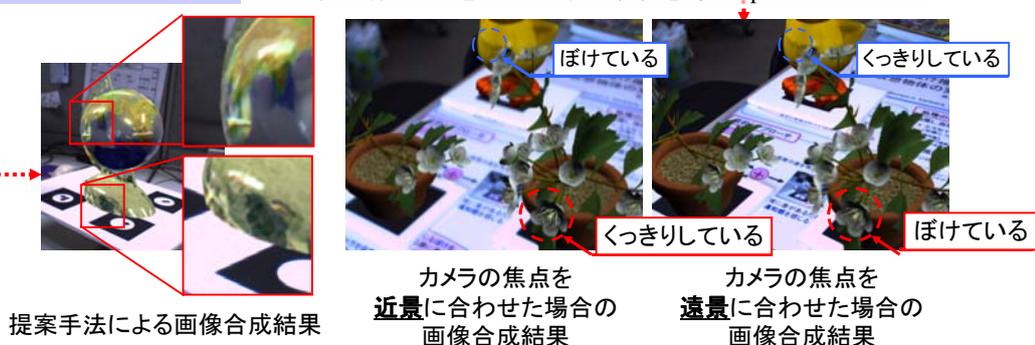


## システム構成



## 画像合成結果

実画像とぼけを再現した仮想物体を 約10fpsで合成可能



# ハイダイナミックレンジ画像からの光源環境推定による 拡張現実感のための光学的整合性の実現

視覚情報メディア講座 仁科 勇作, 奥村 文洋, 神原 誠之, 横矢 直和

## 研究背景

### 光学的整合性の実現

写実的な仮想物体の描画が要求される拡張現実感では、光学的整合性の実現が望まれ、実環境と仮想環境の間の

- ・影・陰影
- ・写り込み

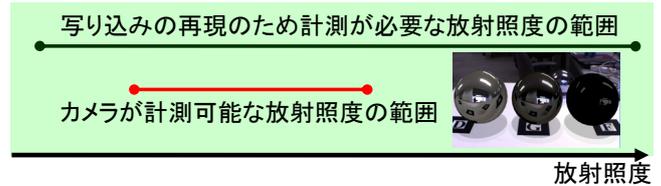
の矛盾を解決する必要がある



製品設計への応用例

### 光学的整合性問題の解決に関する従来研究の問題点

光源環境を推定するカメラのダイナミックレンジ不足による、仮想物体の写り込みにおける写実性の低下を考慮していない



## 目的とアプローチ

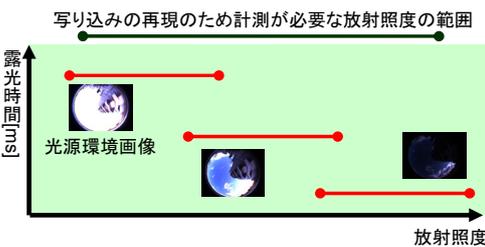
目的 動的な光源環境に対し、カメラのダイナミックレンジ不足の問題を解決することで、写実的な写り込みを再現可能なARを実現

### アプローチ



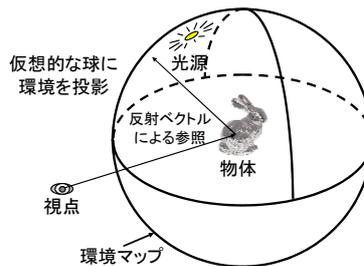
### HDR画像生成

露光時間を段階的に変更して、異なる放射照度の範囲を計測し、HDR光源環境画像を合成



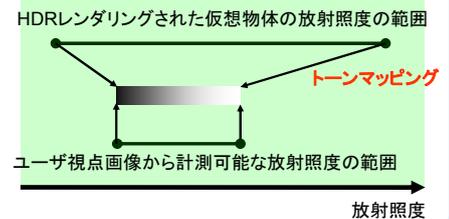
### HDRレンダリング

HDR光源環境画像を環境マッピング

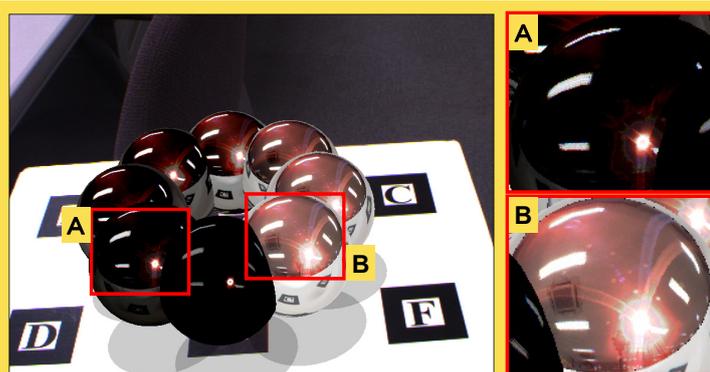


### トーンマッピング

HDRレンダリングされた仮想物体とユーザ視点画像のダイナミックレンジを一致させる



## 画像合成結果



HDR光源環境画像からの  
光源環境推定によるAR合成画像



光源環境推定に用いるカメラのダイナミックレンジ  
不足を考慮しない場合のAR合成画像

# 装着した3次元磁気センサを用いた 歩き・走り状態に対応したユーザの自己位置推定

山中 一樹, 神原 誠之, 横矢 直和

URL: <http://yokoya.naist.jp/>

## 研究背景

モバイル端末を持ったユーザへの位置依存情報提示

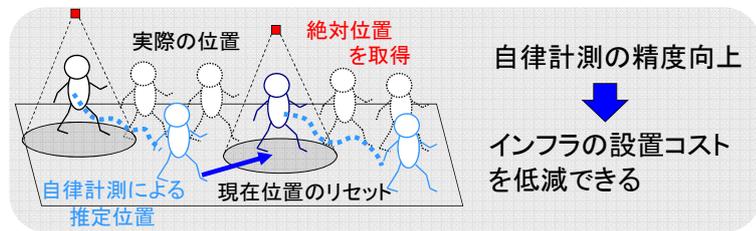
- 応用例**
- ヒューマンナビゲーション
  - エンターテイメント

ユーザの **姿勢**: コンパス・ジャイロなどで取得可  
**位置**: 多様な環境で取得できる手法がない

## ユーザの位置推定手法

屋外環境: GPSを用いることが一般的  
屋内環境: **インフラによる絶対位置の取得**  
自律計測による**相対移動量の推定** } 組み合わせて推定

- 絶対位置を取得可能
- 装着センサのみで推定可能
- × 利用範囲に比例して設置コスト増
- × 蓄積誤差の発生
- × ユーザの動作を制限



ユーザの動きに制限の無い自律計測手法が望まれる

## 目的

### 走りにも対応可能な相対移動量計測手法の開発

#### アプローチ

ユーザの歩数と移動動作を計測

#### 計測対象

- 腰と各足の相対位置関係(3次元磁気センサ)
- 腰の絶対姿勢(姿勢センサ)
- 足の接地タイミング(押しボタンスイッチ)

#### 相対移動量

- 左右の踵の相対位置
- 両足が宙に浮く期間では腰は放物運動しているとして腰の移動量を計算

## 提案システムの機器構成



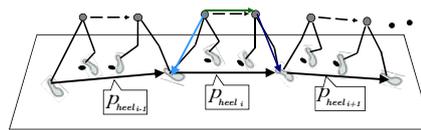
\* [ ]: センサの装着箇所

## 自己位置推定手法

### ◆推定手順

1. 接地足の判定
2. ステップベクトルの計算
3. ステップベクトルの加算

$$p_{heel} = [R_w(T_{f1} + R_f T_{f2}) - R_{w(-\Delta t)}(T_{b1} + R_b T_{b2})] \cdot L$$



### ◆推定移動量のキャリブレーション

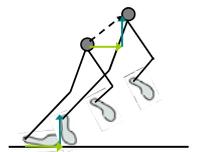
問題点: ふみきり足の踵が浮き上がる  
解決策: 踵の浮き上がり量を推定して補正

- 定められた距離を一定の速度で数回走る
- 各回ごとに速度を変化させて走る

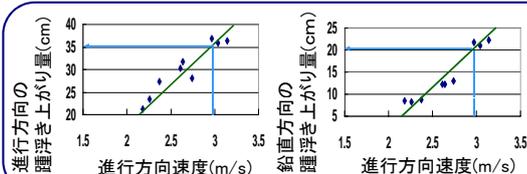
仮定: 両足が宙に浮く期間では腰は放物運動する

両足が宙に浮く期間の移動量

$$L = \begin{bmatrix} v_x \Delta t \\ v_y \Delta t \\ v_z \Delta t - \frac{1}{2} g \Delta t^2 \end{bmatrix}$$



学習 → 速度と誤差の傾向



進行方向速度に応じた踵浮き上がり量を腰の位置に加算

## 自己位置推定結果

### ●実験経路

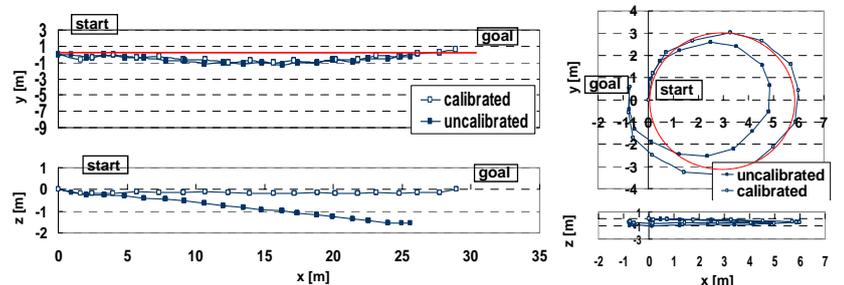
30mの直線走り  
直径6mの円状走り

### ●評価指標

位置推定誤差率  
終点の位置推定誤差 / 総移動量 [%]

### 位置推定誤差率

- ◆ 直線走り 3.4[%]
- ◆ 円状走り 4.9[%]



# 被写体距離を考慮したImage-Based Rendering による広域屋外環境のステレオ画像生成

視覚情報メディア講座 堀 磨伊也, 神原 誠之, 横矢 直和

## 背景

遠隔地の情景を提示し、その場にいる感覚を与えるテレプレゼンスに関する研究が盛ん

臨場感豊かに情景を提示することが望まれる

## 目的とアプローチ

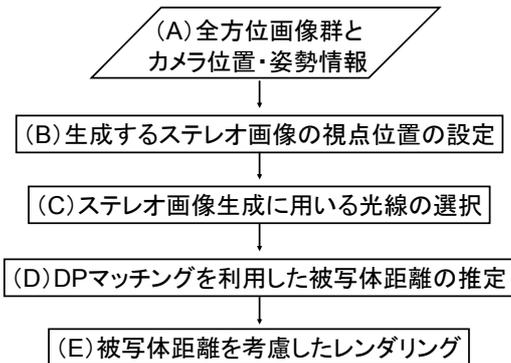
### 目的

自由な見回しと立体視を可能とする  
屋外環境のテレプレゼンスシステムの構築

### アプローチ

- 広域屋外環境を対象とするために  
全方位画像群とカメラ位置・姿勢情報を同時に取得
- 立体視を可能とするために  
被写体距離を考慮してImage-Based Renderingにより両眼ステレオ画像を生成

## 処理手順

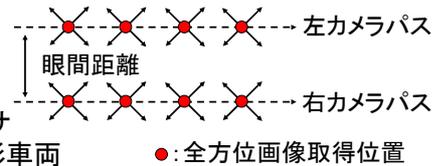


## 提案手法の概要

### (A) 全方位画像群とカメラ位置・姿勢情報の取得(屋外の自由経路)



全方位カメラ・位置姿勢センサ (RTK-GPS+INS) を搭載した撮影車両



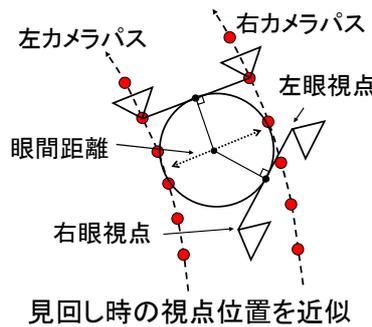
左カメラ画像

右カメラ画像



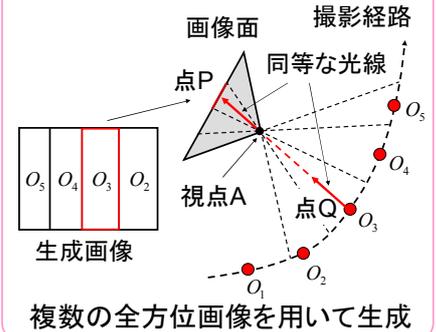
全方位パノラマ画像 (2048 × 1024 画素)

### (B) 視点位置の設定(見回し時)



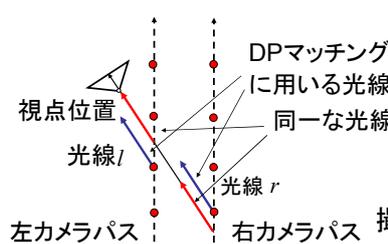
見回し時の視点位置を近似

### (C) 生成に用いる光線の選択

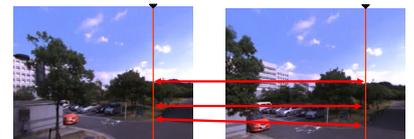


複数の全方位画像を用いて生成

### (D) DPマッチングを利用した被写体距離の推定



左カメラパス 右カメラパス



DPマッチングでエッジの対応付け  
撮影位置情報を用いて被写体距離推定

## 生成されたステレオ画像

360度すべての視線方向の画像を生成



被写体距離考慮なし (従来手法) 被写体距離を考慮 (提案手法)



左眼画像

右眼画像

## テレプレゼンスシステム



(シャッター式眼鏡+ゲームコントローラ)

立体視と自由な視点移動・視線変更が可能

# テクスチャの明度変化と局所性を考慮した パターン類似度に基づく エネルギー最小化による画像修復

奈良先端科学技術大学院大学 視覚情報メディア講座  
河合 紀彦, 佐藤 智和, 横矢 直和

## 画像修復 (Image Inpainting)

写真についての傷や意図せず写りこんでしまった物体などの画像内の不要部分を取り除き、取り除かれた領域(欠損領域)を自動的に違和感なく修復

## アプローチ

概念

欠損領域内の全てのテクスチャが  
データ領域のテクスチャと類似



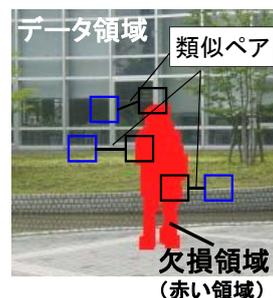
違和感の少ない画像ができるのではないかと

**パターン類似度に基づいたエネルギー関数を定義し、それを最小化することで画像修復**

しかし

単純なパターン類似度のみによるエネルギー関数では、

画像内の 証明条件等の変化による明度変化  
透視投影等によるパターン構造の変化 } に対応できず不自然な画像が生成される



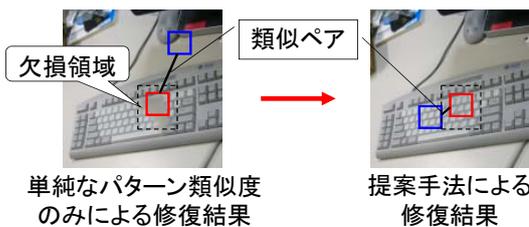
## パターン類似度によるエネルギー関数の拡張

### 拡張1: 輝度補正係数の導入

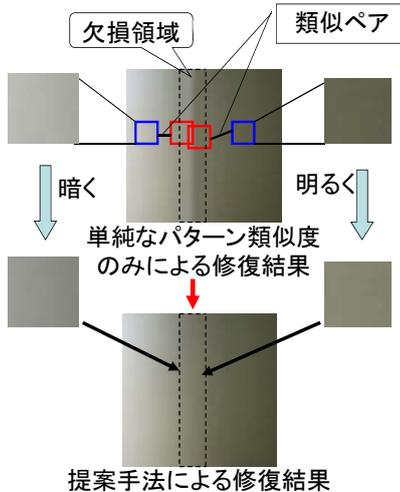
テクスチャの線形的な明度変化を許容  
不自然な明度変化による違和感の軽減

### 拡張2: 画像の局所性を考慮した項の追加

近傍テクスチャへのマッチングコストを低下  
不適切なマッチングによる不鮮明なテクスチャ生成の抑制



ぼけを防止し、細かいテクスチャを再現



滑らかな明度変化を再現

## 処理手順

画像1枚を入力

対象領域(欠損領域)を手動で入力

エネルギー最小化による画素の更新

エネルギーが収束したか?

No

Yes  
画像1枚を出力

## 実験結果



原画像

欠損画像

修復画像



原画像

欠損画像

修復画像



# 局所形状の類似度に基づく エネルギー最小化による三次元欠損修復

奈良先端科学技術大学院大学 視覚情報メディア講座  
河合 紀彦, 佐藤 智和, 横矢 直和

## 三次元欠損修復

実物体に対する計測を行う際のオクルージョン等による計測漏れによって生じる三次元モデル内の欠損領域を自動的に違和感なく修復

## アプローチ

概念 同一物体内に欠損領域内の形状と類似する形状が存在する可能性が大きい → 類似した局所的な形状を合成することで、違和感なく修復できるのでは？

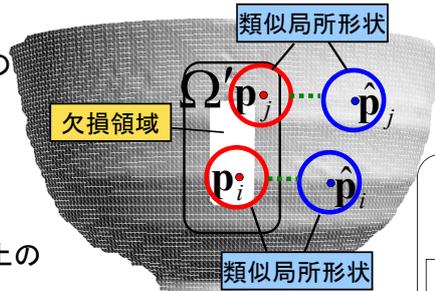
局所的な形状の類似度に基づくエネルギー関数を定義し、それを最小化することで欠損領域を修復

### エネルギー関数の定義

欠損領域を含む領域  $\Omega'$  内の頂点  $\mathbf{p}_i$  とそれ以外の領域の頂点  $\hat{\mathbf{p}}_i$  の周辺の局所形状の類似度SSDの重み付き和

$$E = \frac{\sum_{\mathbf{p}_i \in \Omega'} w_{\mathbf{p}_i} SSD(\mathbf{p}_i, \hat{\mathbf{p}}_i)}{\sum_{\mathbf{p}_i \in \Omega'} w_{\mathbf{p}_i}}$$

エネルギーが小さい → 欠損領域内の形状が同一物体上のどこかの形状と類似している

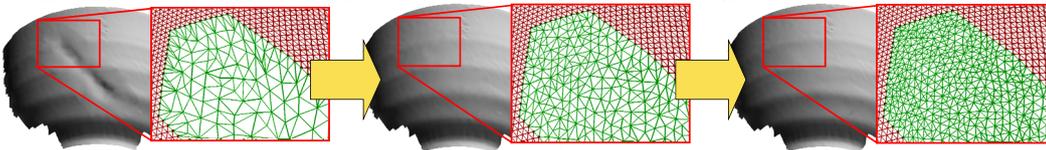


### エネルギー関数の最小化手法

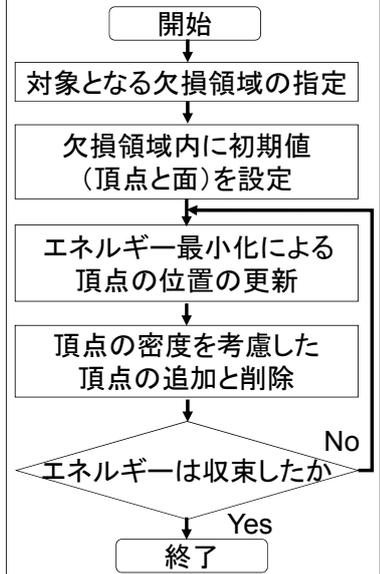
処理(i)と処理(ii)をエネルギーが収束するまで繰り返すことで、欠損領域を修復

- (i) 欠損領域内の全ての頂点の位置を固定し、欠損領域内の頂点  $\mathbf{p}_i$  と最も類似した(Eが最小となる)形状を持つ頂点  $\hat{\mathbf{p}}_i$  の探索
- (ii) (i)で求まる欠損領域内の頂点と類似した頂点の位置を固定し、類似した頂点周辺の形状を用いて、欠損領域内の頂点  $\mathbf{p}_i$  の位置を更新

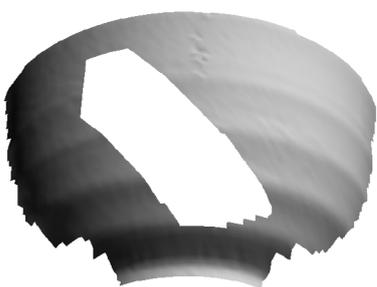
頂点の密度を上げながら上記のエネルギー最小化処理を反復



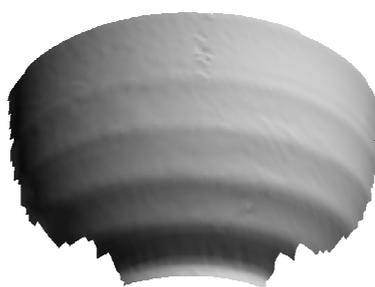
### 処理手順



## 実験結果



欠損をもつモデル



修復したモデル



欠損をもつモデル



修復したモデル