

浮草研究者と人工知能研究

文化遺産のモデル化と人まねロボット

池内克史

米国マイクロソフト・研究開発本部

応用ロボット研究チーム

人は誰でも生まれつき知ることを求める

— アリストテレス 形而上学

- 旅人： 新しい場所を訪問、その地を知る喜び

- 研究者： 新しいテーマを彷徨い、その問題を知る喜び

京大工機械
東大院情報

電総研(5)

MIT(3)

CMU(11)

東大(18)

Microsoft(7)

- 京大から東大: その学年は東大入試中止で、東大の卒業生がない 両方受かった 東京の方が実家から遠かった
- 東大からMIT: アメリカで研究してみたかった
- MITから電総研: アメリカでDARPA(アメリカ)のために研究をするよりも、日本で日本のために研究するべき
- 電総研からCMU: まだまだ、自分は未熟で、日本のために役にたつような人材ではないと痛感
- CMUから東大: そろそろ日本に戻って、日本の研究者を育てよう
- 東大からMicrosoft: 定年になった

ロボットビジョン

ビンピッキング



物体認識プログラムの自動生成



人間行動観察学習ロボット



コンピュータビジョン

明るさ解析

- 単眼画像解析
- 照度差ステレオ
- 透明物体解析

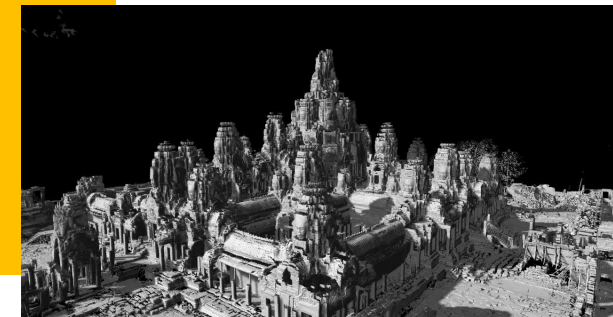
反射率解析



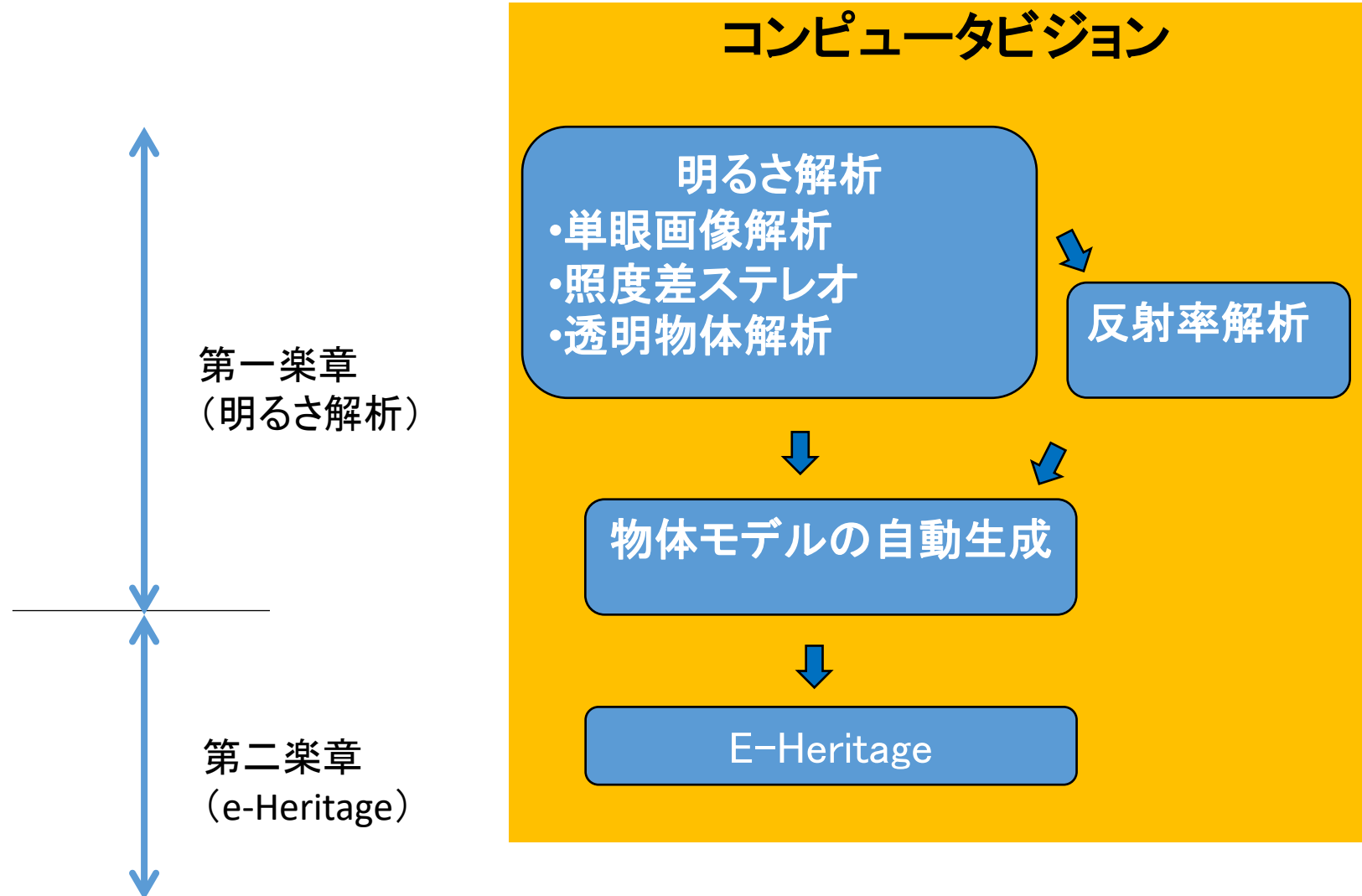
物体モデルの自動生成

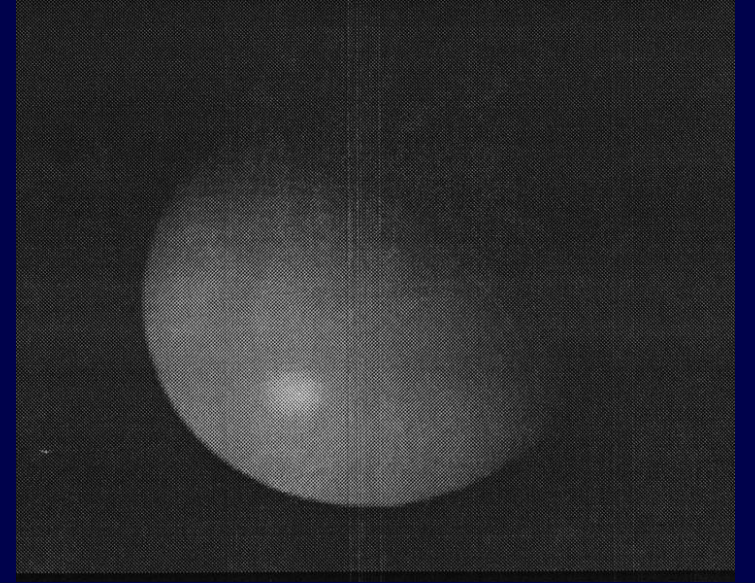


E-Heritage



まずは前半のコンピュータビジョンでの放浪





第一楽章 (Andante) 明るさ解析

マサチューセッツ工科大学・人工知能研究所
工業技術院・電子技術総合研究所

なぜ、明るさ解析？

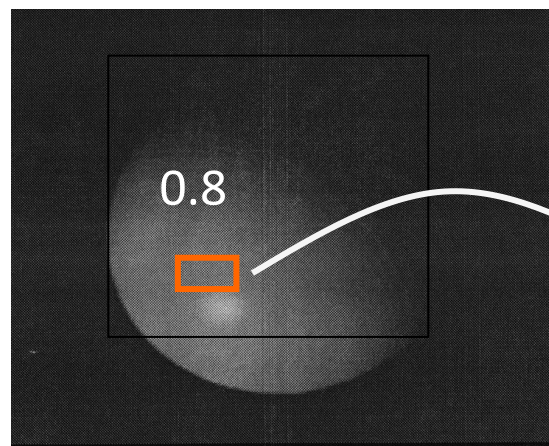
- 東大での博士論文は、「人間の文字(漢字)認識機構のモデル化」
- 博士修了後、なんとなくMIT行ってみたかった
- 東大に来たMinskyにタクシーと一緒に乗り込み、頼み込む(白井さんのおかげ)
- Horn と Marr がその当時MITAIラボにいたが、所長のWinstonは、Hornを勧める
- たぶん、DARPAのファンディングの影響、その当時、Image Understanding Project が走っていた
- 辺境の地日本から、なんとなく中心のMITに来て、その中で目立つ必要があった
- Horn のグループはみんな明るさ解析

研究環境(偶然)がテーマを決める

明るさ解析の定義

- ・ 絵画を見た時に、1枚の絵にかかわらず奥行感がある
- ・ 人間は、陰影情報から、面の傾きを得ている

明るさ E
(1自由度)



(p, q) 傾き
(2自由度)

→ 明るさは、一変数、面の傾きは、2つの変数
明るさの観測から、未知数2個の1本の方程式

$$E = R(p, q)$$

もう一つ方程式が必要！

滑らかさ拘束： 隣り合った点は、おなじような傾きを持つ

$$\left(\frac{\partial p}{\partial x}\right)^2 + \left(\frac{\partial p}{\partial y}\right)^2 + \left(\frac{\partial q}{\partial x}\right)^2 + \left(\frac{\partial q}{\partial y}\right)^2 = 0$$

p and q -- surface orientations

変分法

- 観測された明るさ E と傾きから予想される明るさ $R(p,q)$ が一致

$$F = \int (E(x, y) - R(p(x, y), q(x, y)))^2$$

- 隣り合う点は同じような傾き

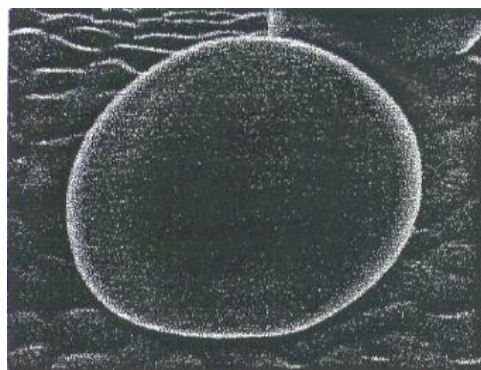
$$+ \lambda \left\{ \left(\frac{\partial p}{\partial x} \right)^2 + \left(\frac{\partial p}{\partial y} \right)^2 + \left(\frac{\partial q}{\partial x} \right)^2 + \left(\frac{\partial q}{\partial y} \right)^2 \right\} dx dy \quad \rightarrow \min$$

- 変分法により繰り返し公式に変化できる

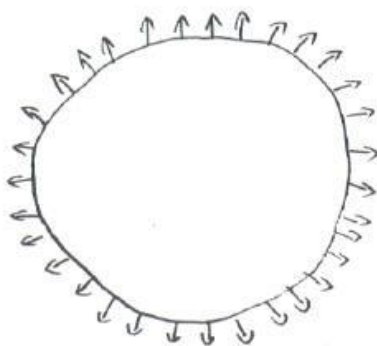
$$p^{n+1}(x, y) = 1/4 \{ p^n(x-1, y) \dots \} + \lambda \{ \dots \}$$

$$q^{n+1}(x, y) = 1/4 \{ q^n(x-1, y) \dots \} + \lambda \{ \dots \}$$

滑らかさ拘束による明るさ解析



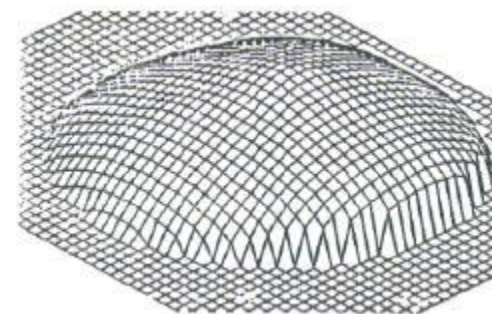
1枚の濃淡画像



境界条件としての輪郭線



明るさ一致と
滑らかさ拘束



形状

N 番目の解

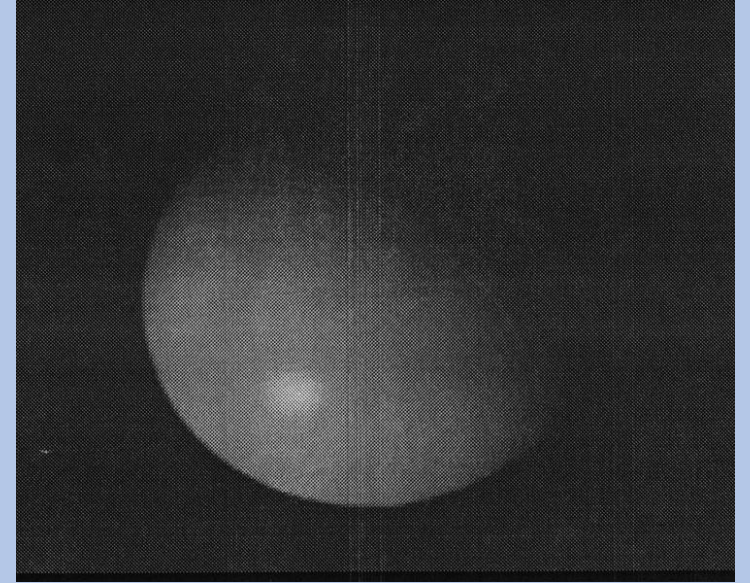
$N+1$ 番目の解

繰り返し法

実のところ、、、

- ぐちゃぐちゃハッキングしている途中
- ある日突然アルゴリズムが収束！
- 理由を調べたところ、隣となるべく同じような傾き
- そこでHornと相談して数式をセットアップし、“滑らかさ拘束”と命名
- その後、1つの定石的な手法となる：
 - Horn-Schunck Optical flow
 - Brooks-Horn Variational approach

⇒ 帰納法的研究の重要性



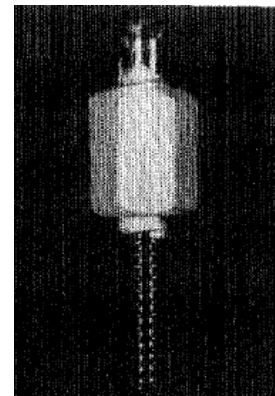
反射率

カーネギーメロン大学・ロボティクス研究所にて

照度差サンプリング



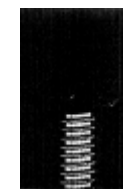
移動光源



形状



乱反射



鏡面反射

反射率パラメータ



Nayar, Ikeuchi, and Kanade *IEEE-TRA* 90
Sato, Nayar, and Ikeuchi *MVA* 90

- 形と反射率が求められる

- それじゃあ実物体から計算機内でのモデルを作ろう！

CMUで主に4つの博士論文

形のモデル化 (幾何パイプライン)

Harry Shum (Microsoft)
Mark Wheeler (Google)



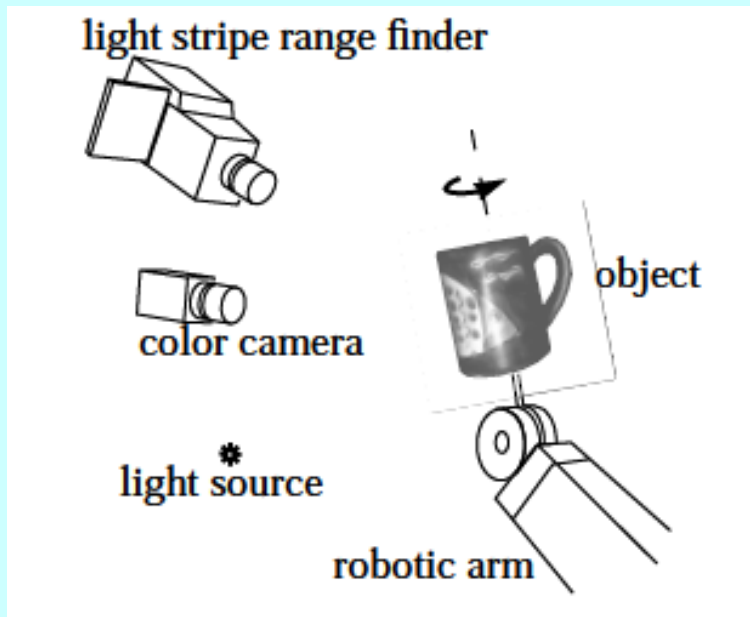
見えのモデル化 (光学パイプライン)

Fred Solomon (Merck)
Yoichi Sato (U Tokyo)



CMUから東大に移るころ (1996年ごろ)

- ・屋内の小さな物体のモデル化はできるようになった



サンプリングセットアップ

幾何パイプライン
光学パイプライン



再構成された見え方 (Y. Sato (東大))

- ・モデリングが面白い！モデルすること自体が重要な分野
⇒ 文化遺産

- ・小さな物体はできた —— 大きな物体で屋外にある文化遺産？

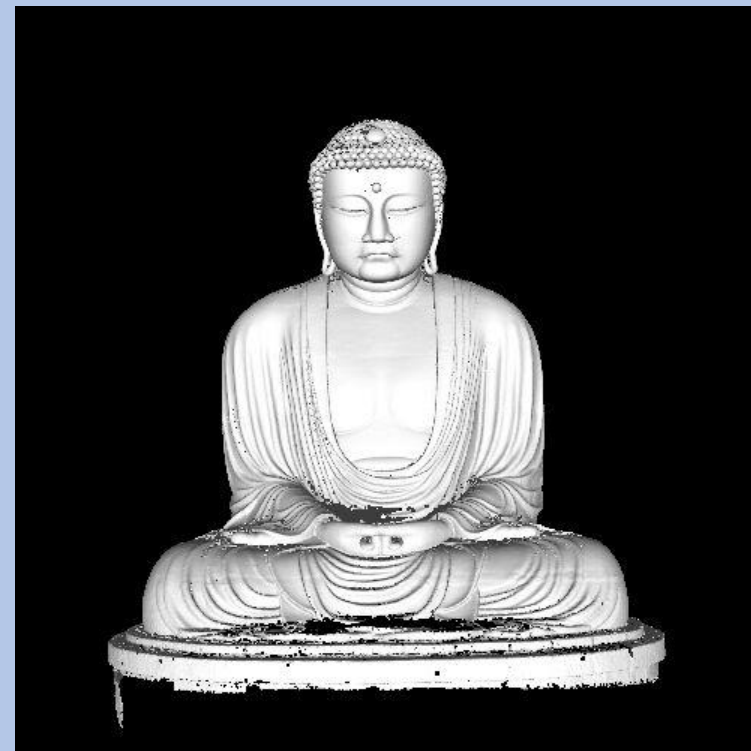
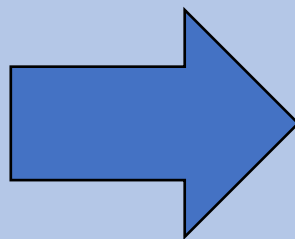
⇒ そうだ！ 鎌倉大仏

- ・坂内先生：



それは面白い
「情報媒介システムプロジェクト」でサポートしましょう！

⇒ **よき理解者を得る**



鎌倉大仏プロジェクト

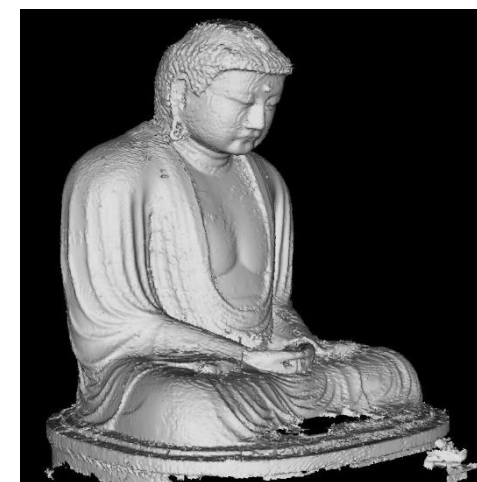
形のモデル化



距離データ収集

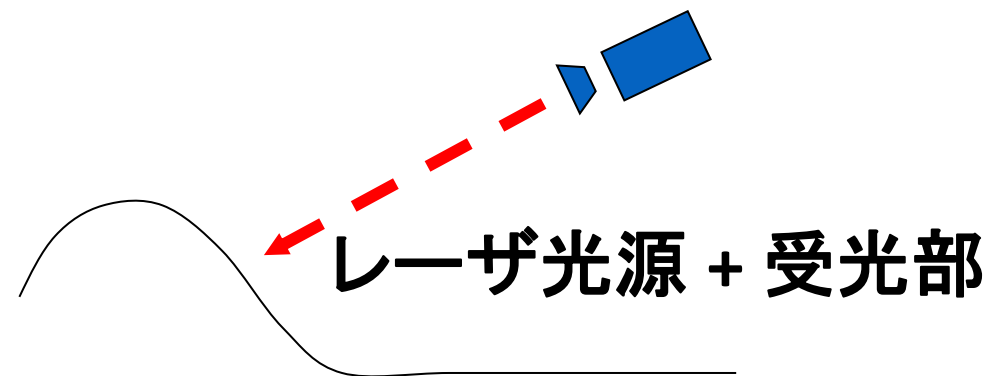


データ相対
位置合わせ



データ統合

データ収集状況 (1999年)

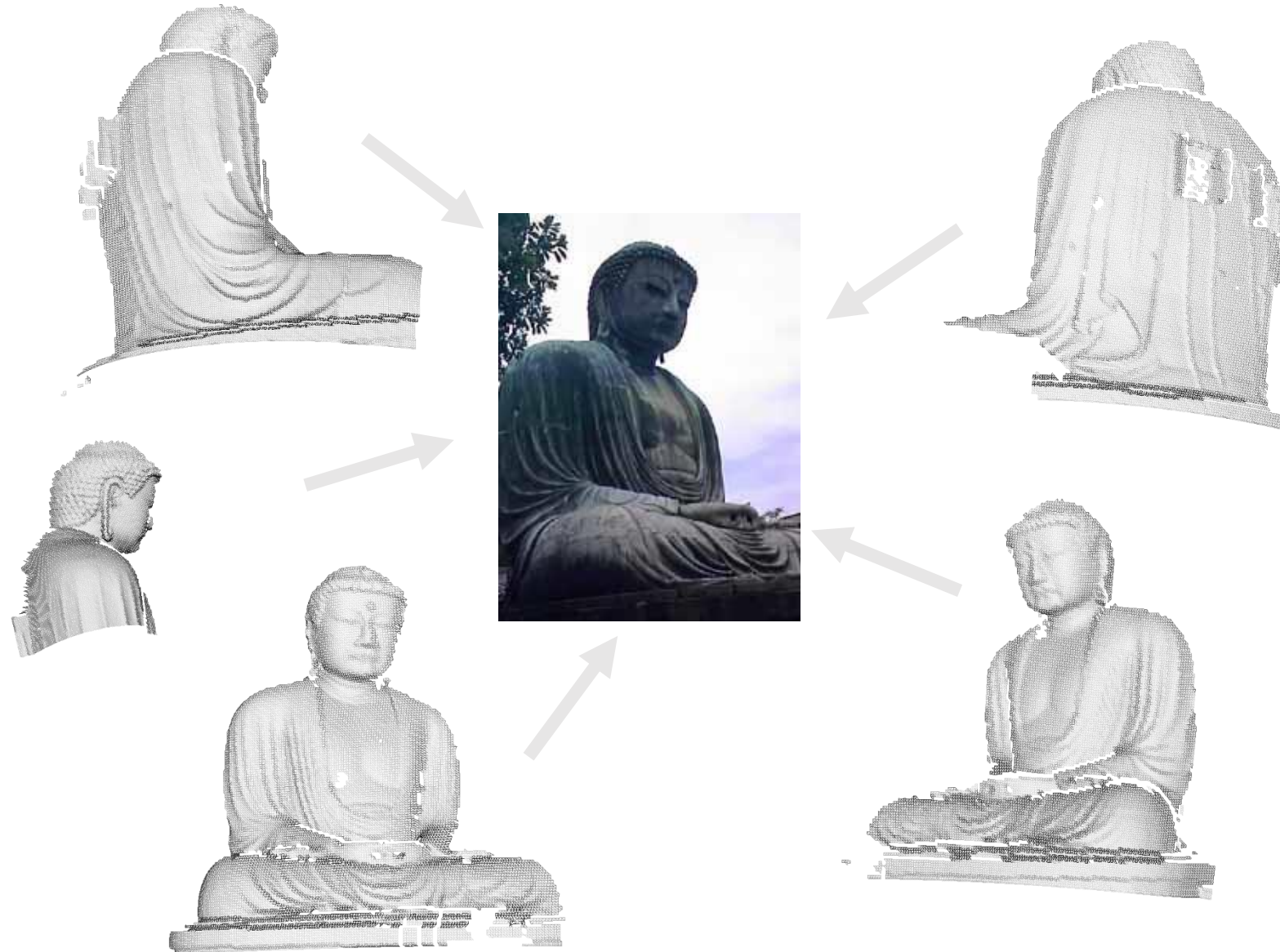


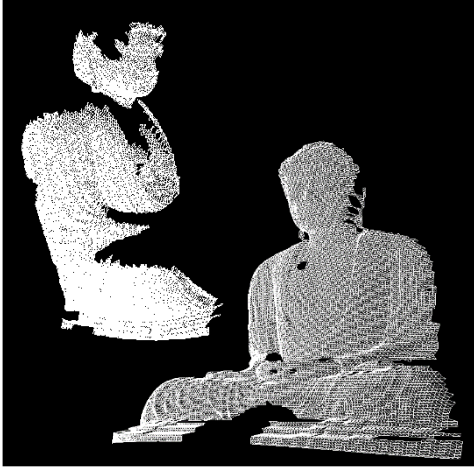
飛行時間 → 距離



CMU時代の学生、Mark
が務めていたスタートアップCYRA社製： 5号機

得られた距離画像





データ収集



データ相対
位置合わせ



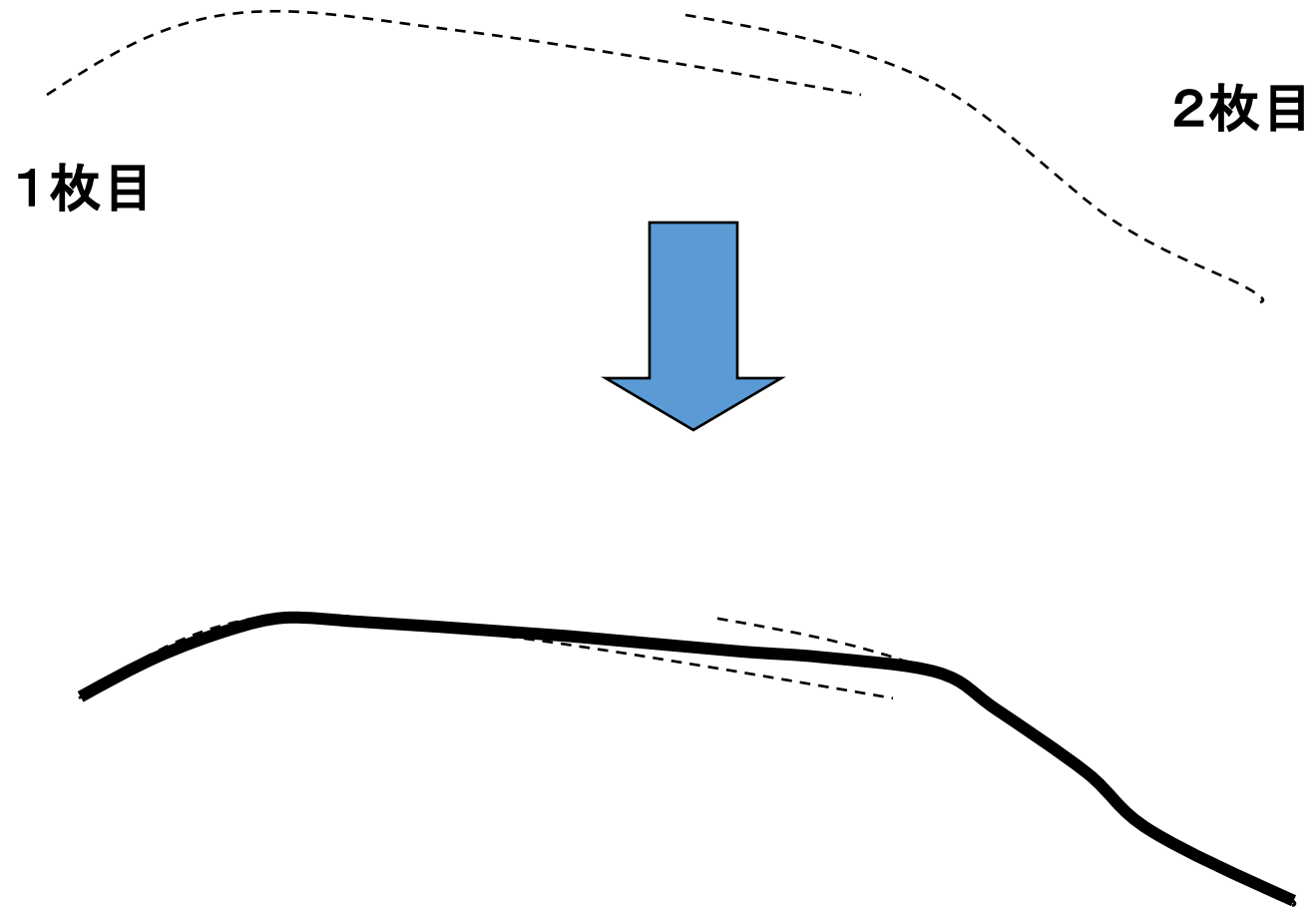
データ統合

位置あわせ

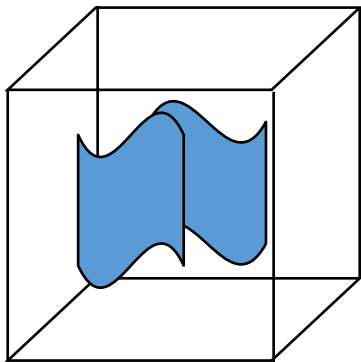


西野 (京大)
倉爪 (九大)

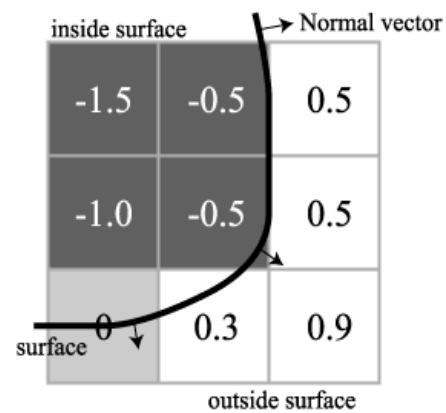
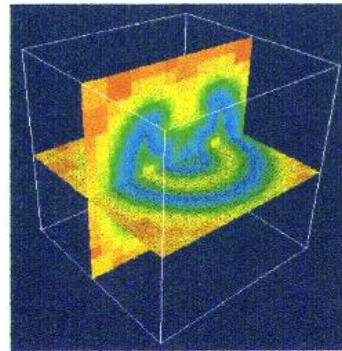
データ統合



データ統合



全ての距離画像を
同一空間に置く



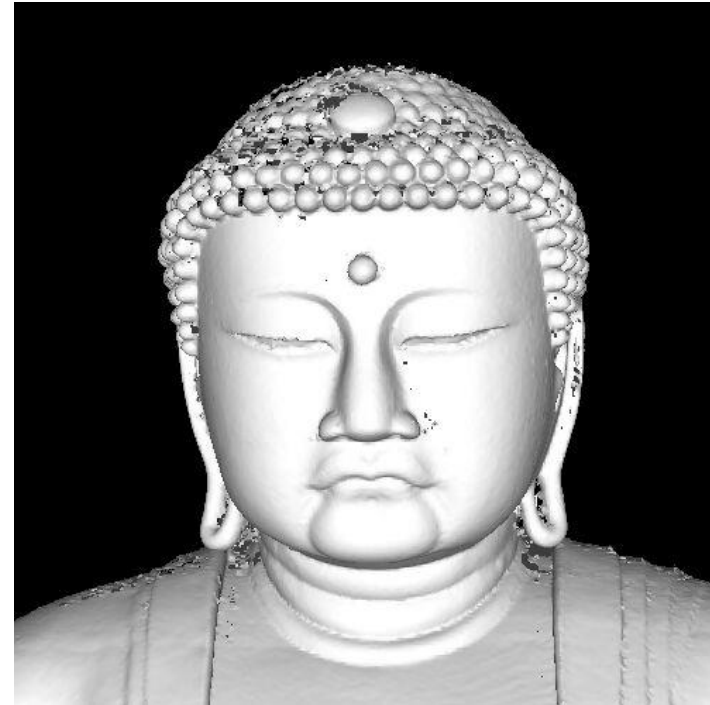
各場所で確から
しさを投票



確からしい場所を
つないでゆく

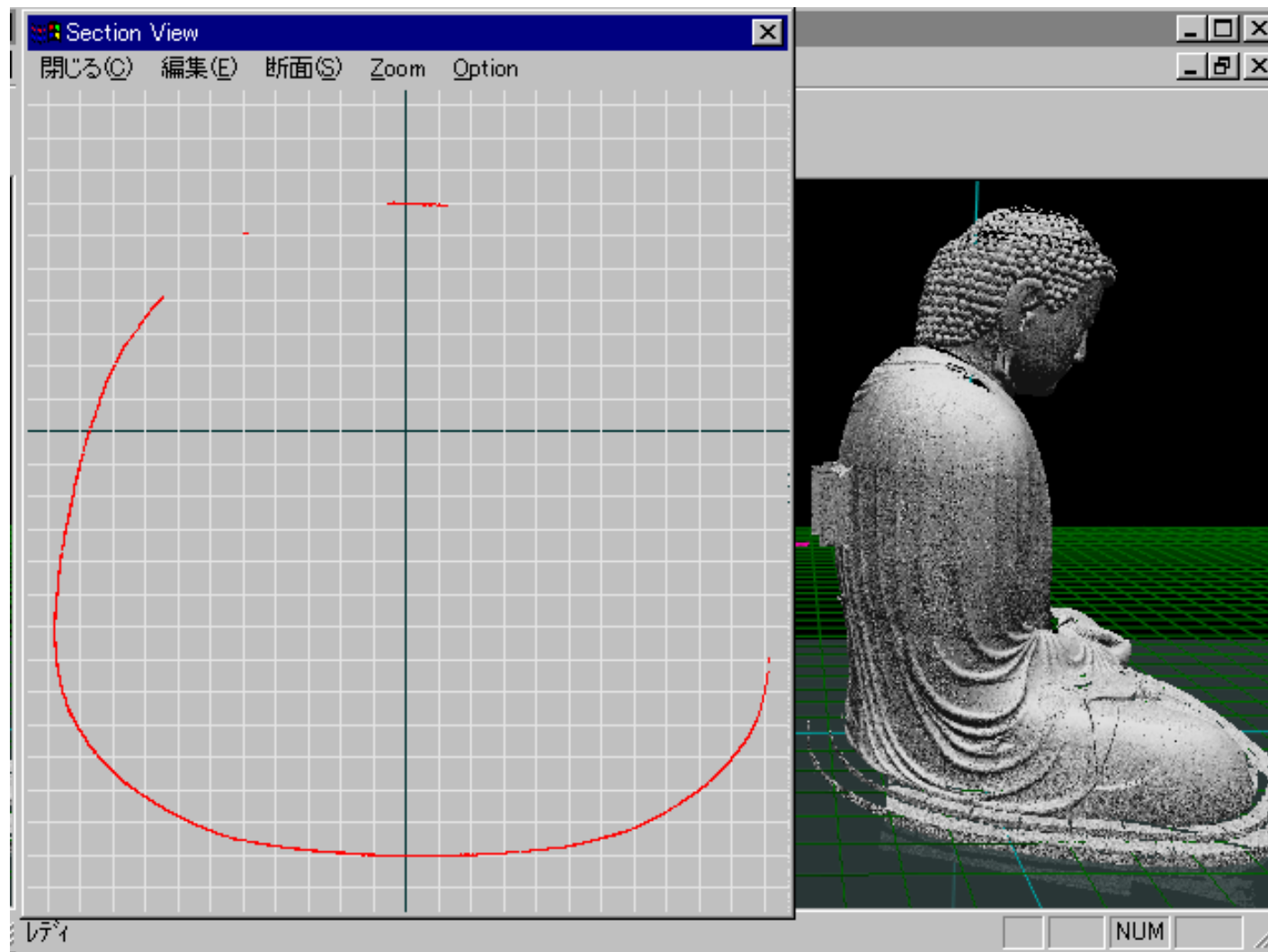
データ統合結果

- 計算時間: 23分 (2000年ごろ)
 - 8 PC (dual PentiumIII 800MHz, 1GB memory)
- 130万点, 550万 ポリゴン



佐川(産総研)
大石(東大)

断面形状



断面形状を見たNHKの記者が
厚さは？と難しいことを質問！

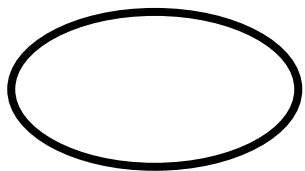
鑄造の厚さ推定

デジタルデータの利活用

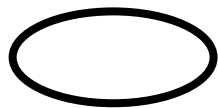
胎内の計測



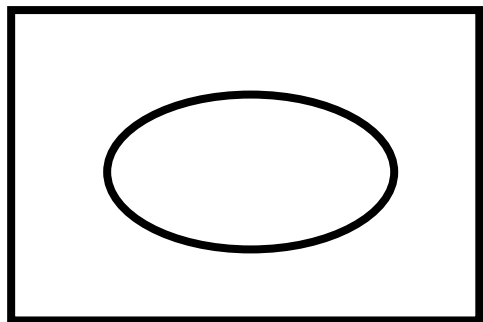
• 2013年



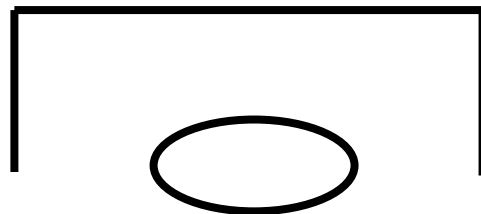
外側データ



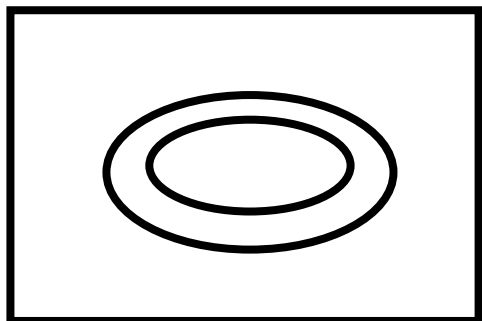
内側データ



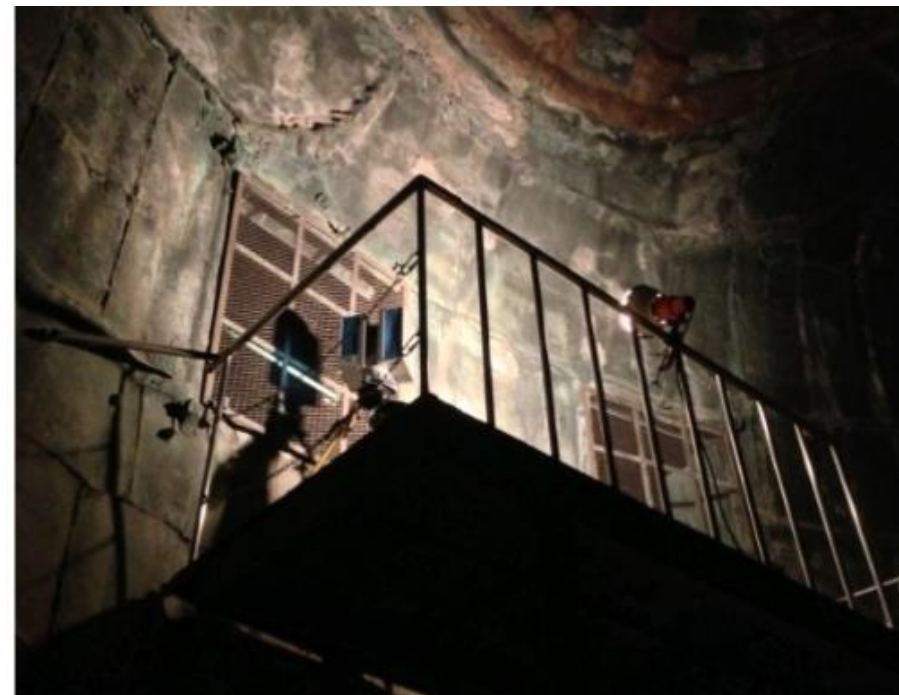
外側＋伽藍



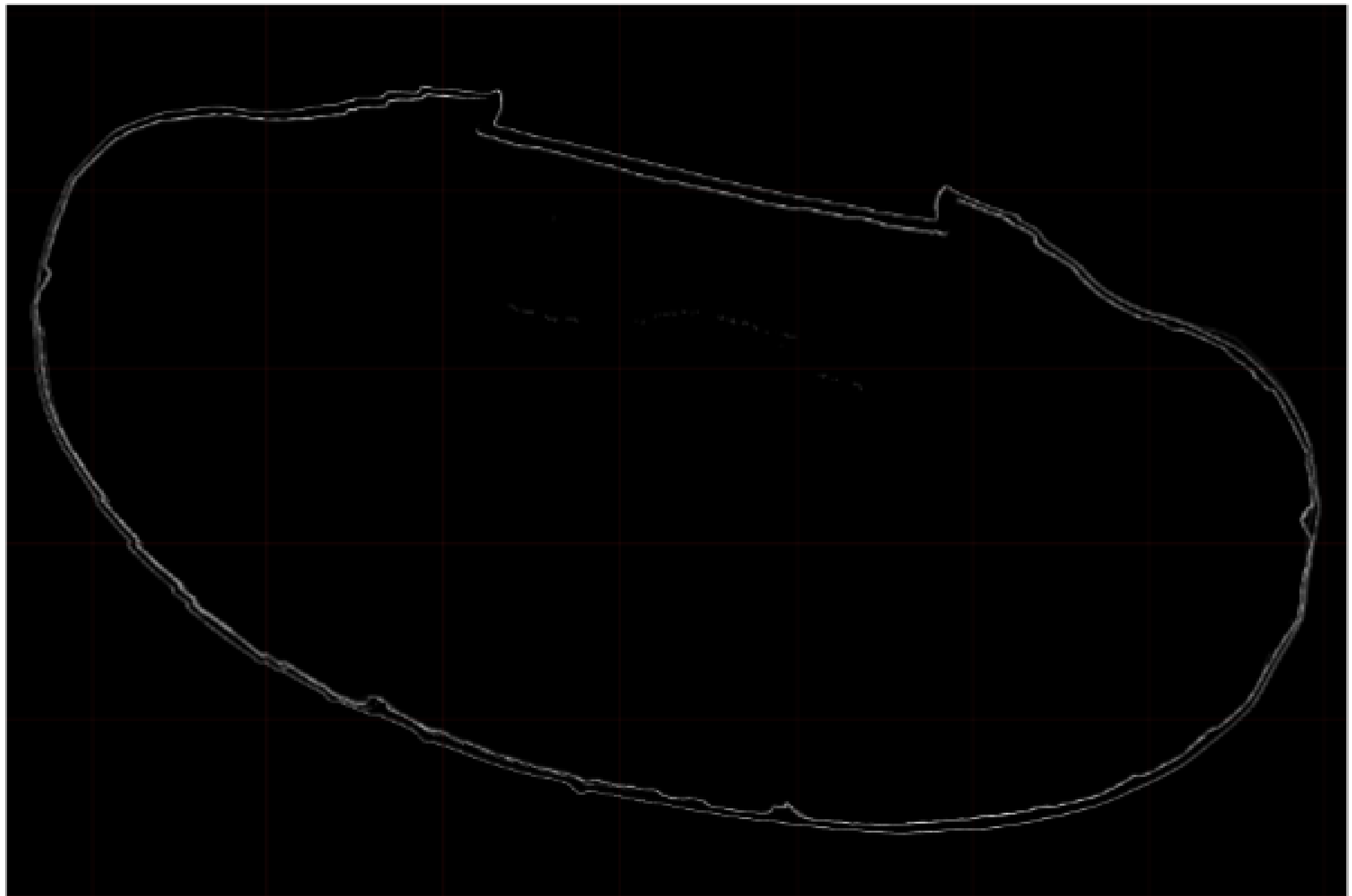
内側＋伽藍一部



伽藍データで重ね合わせ



背開き部から





仮想復元

e-Heritage の利活用

色合いもほしいが、、、



写真の貼り付け
錆や汚れで日々変わる！

もともとは金箔が張られていた

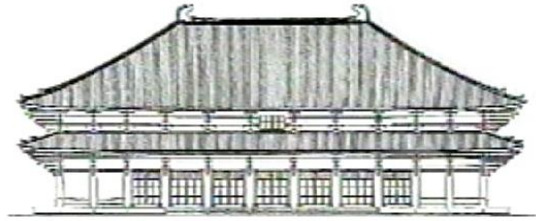


かすかに残っている

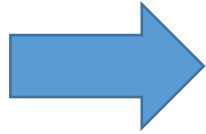
復元結果



昔は大仏殿もあったそうなの、、



東大寺(鎌倉時代)



境内に残る敷石



- 1つでは不十分
- もっと大仏を！

⇒ 継続は力

大仏シリーズ

飛鳥大仏(飛鳥時代)



鎌倉大仏(鎌倉時代)



奈良大仏(江戸時代)



コメント

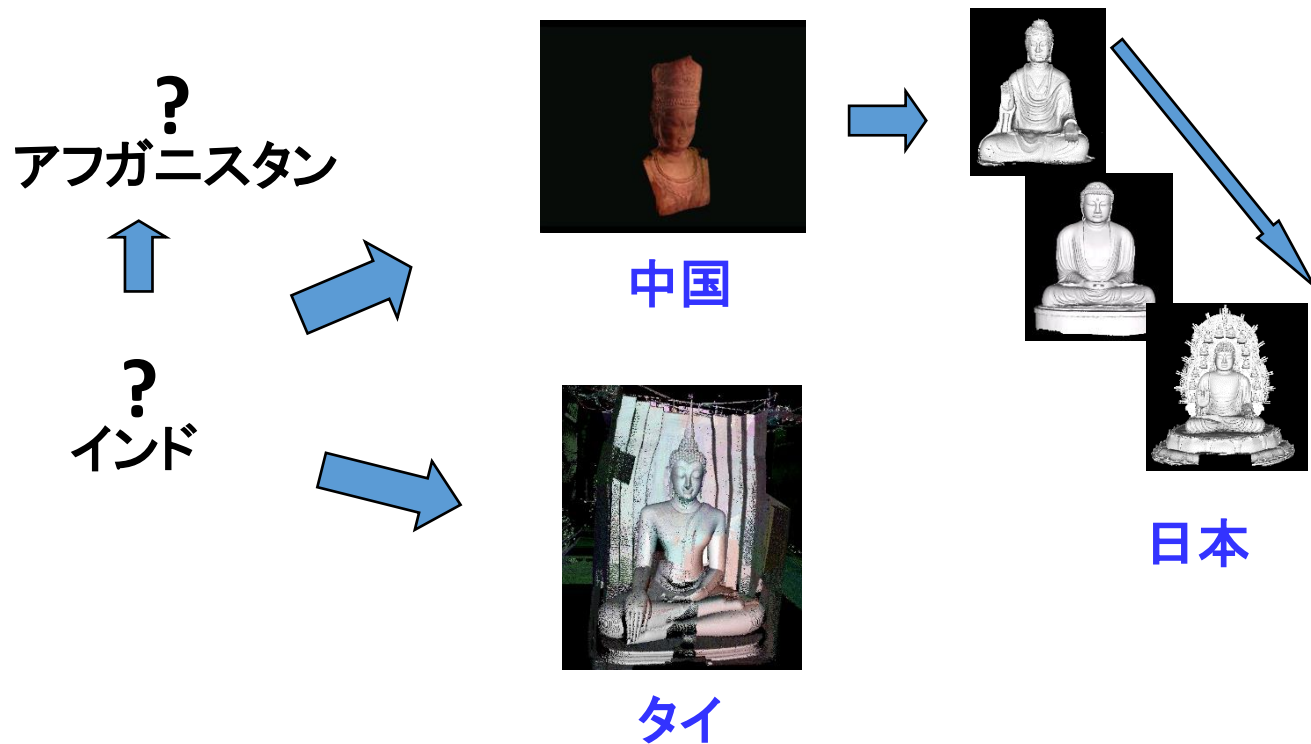
- ・ 日本に大仏、もうない！

⇒ **Think BIG! Do BIG!**



金出先生

世界大仏ライブラリ計画



中国： 断られる

⇒ 後に中国の大学が始める

韓国： 断られる

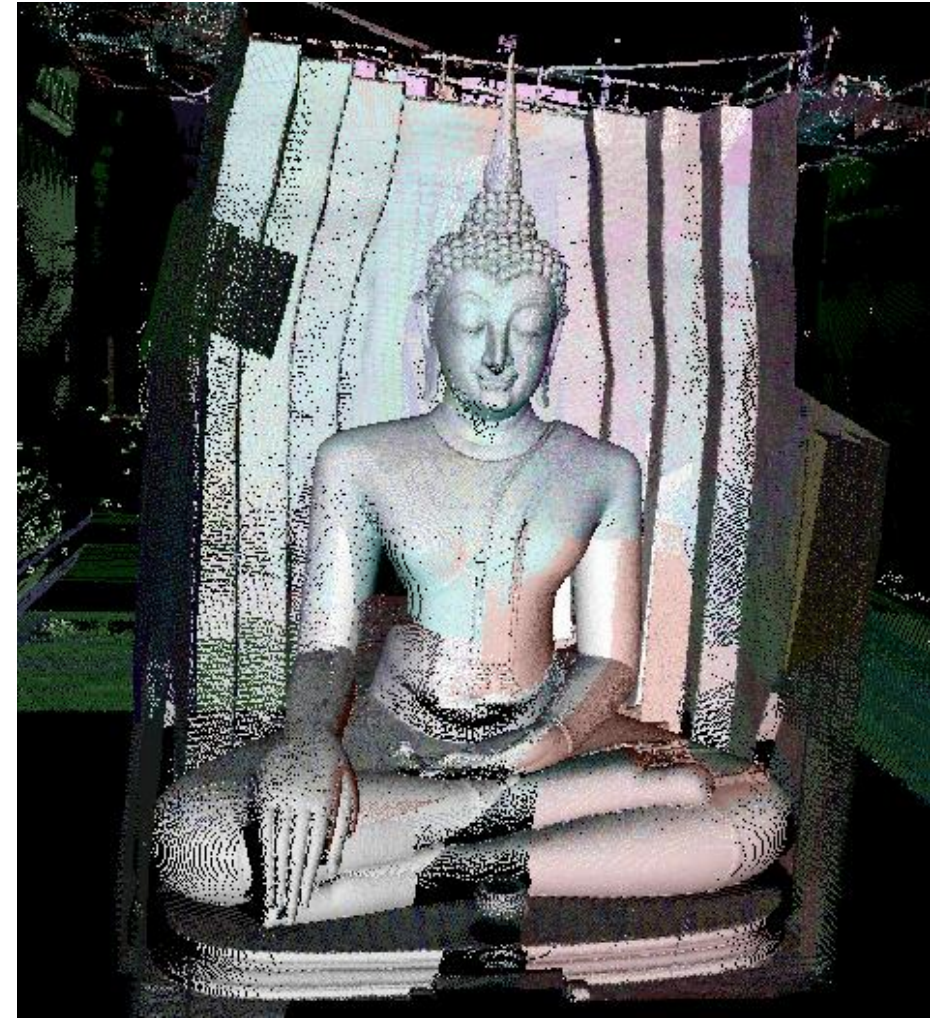
⇒ 後に韓国の大学が始める

タイ： いいですよ！

シーチャム寺院 (スコタイ、タイ)



タイ文化局＋東京文化財研究所
2002年 3月



外壁のため普段はみるこ
とのできないお姿

- ・ スコタイで、タイの文化局の人からランチの時に、

「となりのカンボジアの寺院には、173の大仏の顔があるよ」

⇒ すごい！一気に顔ライブラリができる！

⇒ とにかく動いてみる！

第2楽章 (Allegro)

E-heritage

東京大学・生産技術研究所

東京大学・大学院情報学環



現地に行ってみると、、、顔もよいが！

中央棟崩壊の可能性



→ 現存の間に3Dデータ化

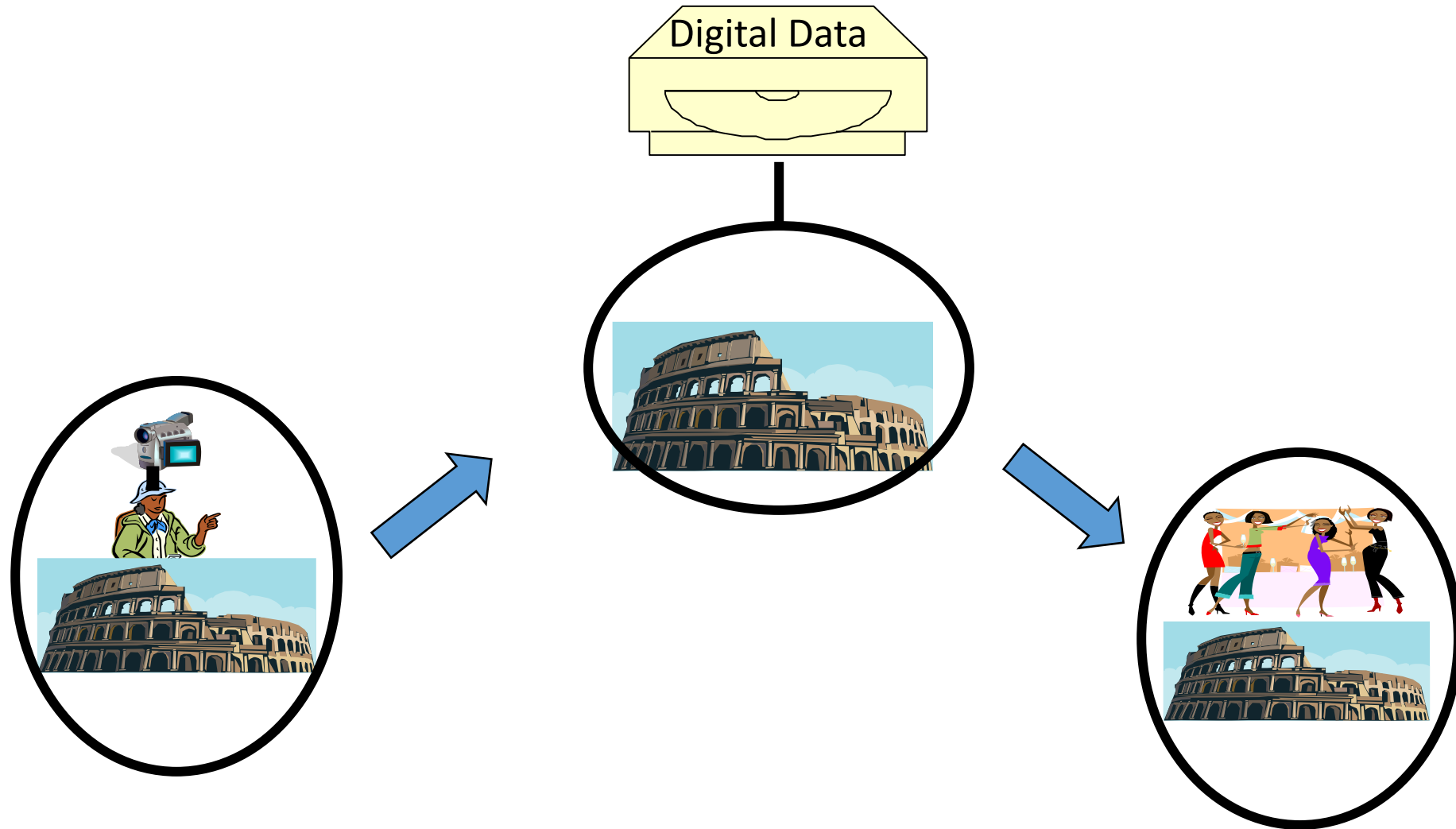
- ・ 大仏だけでなく、寺院全体のモデル化が必要なことを痛感

⇒ できるかなあ???

- ・ そこで、研究の方向性を e-Heritage とせず、JST・クレストに提案

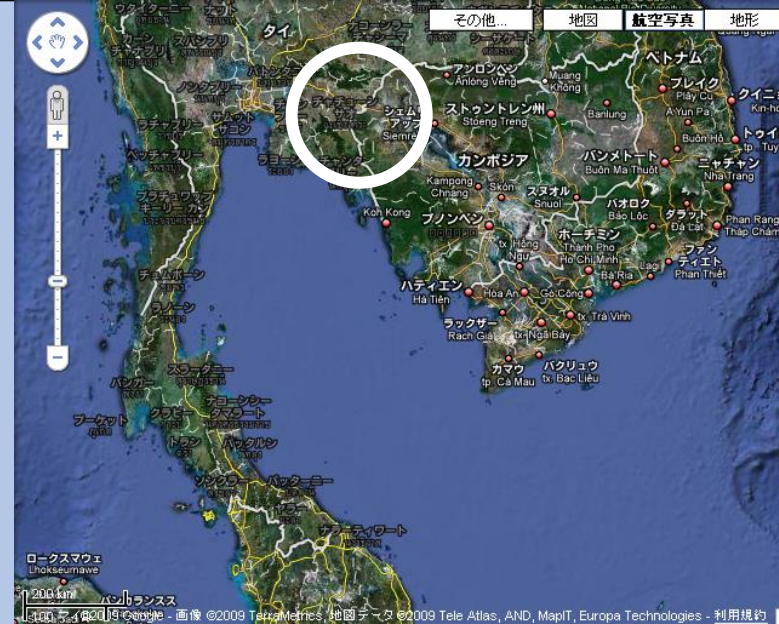
⇒ 必要性は研究のモチベーション

e-Heritage分野の提案



なぜ e-Heritage を推進するのか？

- ・ 非常に高価で保存する価値、鑑賞する価値あり
- ・ なくなる可能性（自然災害、人為行為、後継者難）
- ・ 高価な技術開発の理由づけ（B/Cが高い）
- ・ CS・EEの分野でのフロンティア開発可能



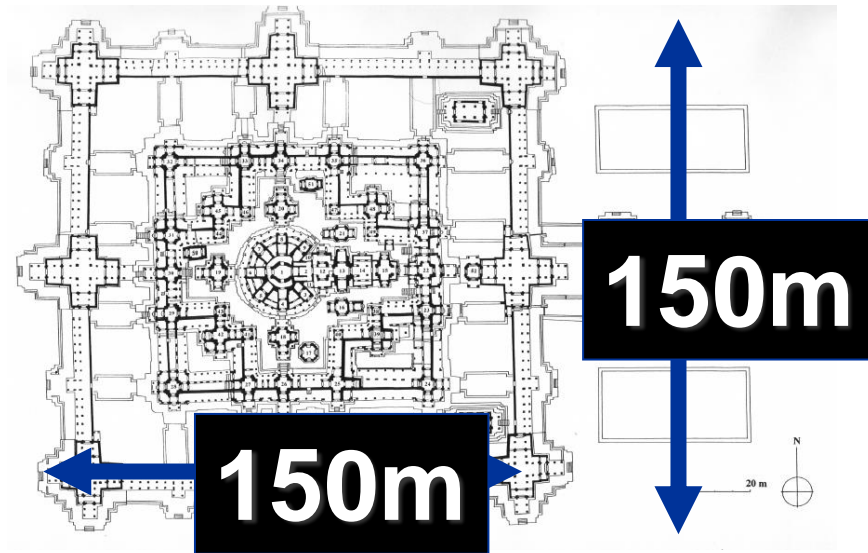
デジタルバイヨンプロジェクト

Ikeuchi et. al. *IJCV* 2007

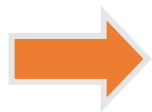
Ikeuchi & Miyazaki Springer 2007

池内・大石 3次元デジタルアーカイブ、東大出版会、2010

しかし、やっぱり、、



⇒ 規模！

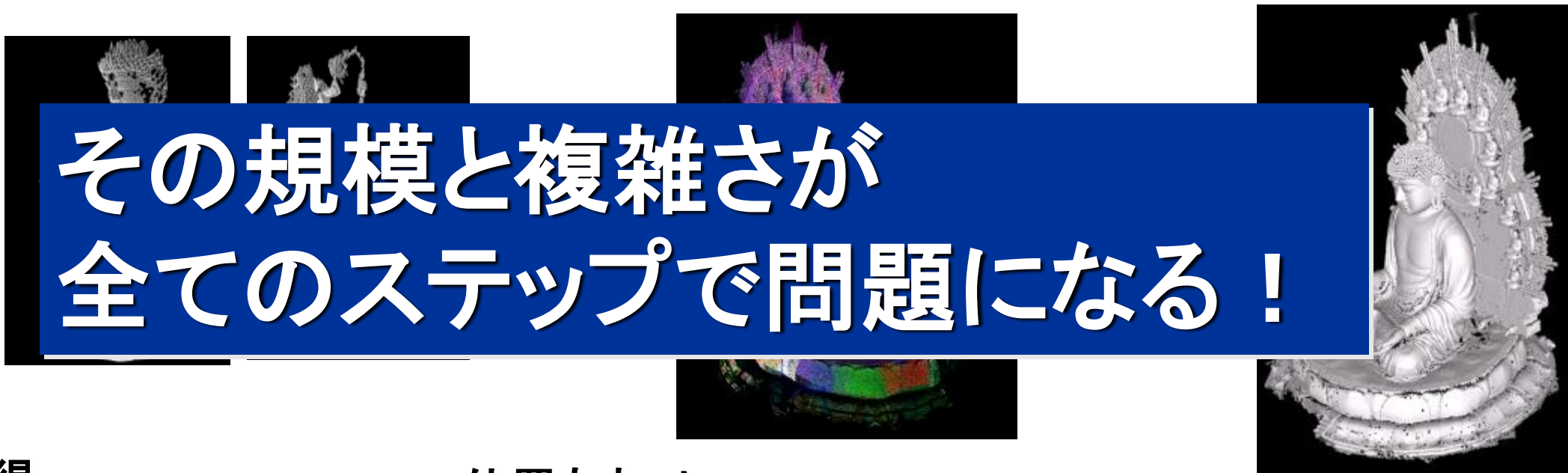


その規模と複雑さが
全てのステップで問題になる！

データ取得

位置あわせ

統合



多数の学生が博士論文テーマとして格闘

センサー設計 ⇒ 博士論文 7本

川崎* (九大)、阪野 (産総研)、小野 (福岡大)、石川" (東大)

森本 (凸版)、増田 (凸版)、Zheng (Huawei)

ソフトウェア設計 ⇒ 博士論文 14本

佐藤い (NII)、西野 (京大)、松下* (阪大)、大石 (東大)、川上 (東工大)

佐川 (産総研)、Tan (NTU)、You (U Amsterdam)、Xin (Pekin U)

山崎 (Apple)、岡本 (スタートアップ)、小林 (富士通)、Huang (Tencent)

鎌倉 (Microsoft)

*坂内先生と指導

"大石先生と指導

(+ 倉爪 (九大)、中澤 (京大)、長谷川 (富士フィルム))

尊顔
回廊内部
回廊壁画

CYRAX(市販)
Z+F(市販)
VIVID(市販)

ペディメント
狹隘部
上空

鏡センサ(池内研開発)
木登りセンサ(池内研開発)
気球センサ(池内研開発)





プレア ヴィヘア

なぜ、プレアヴィヘア？



タイ・カンボジア国境（少しカンボジア内）

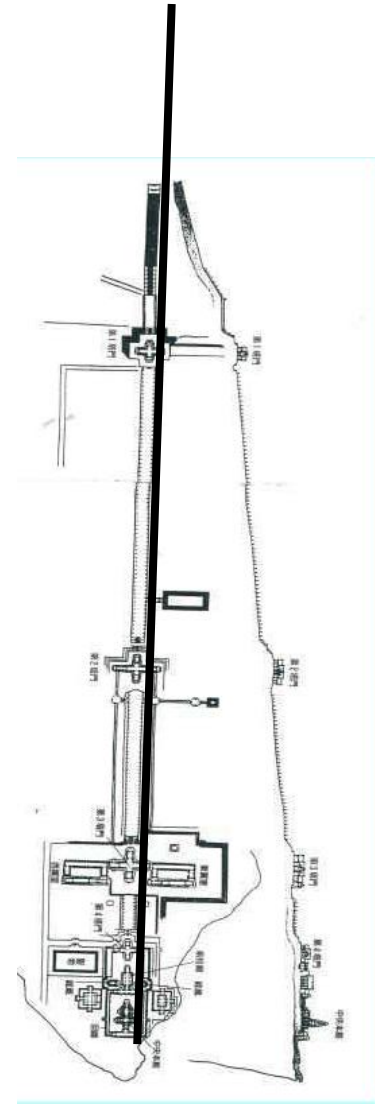
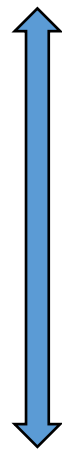
領有権をめぐるたびたび戦争

南北軸

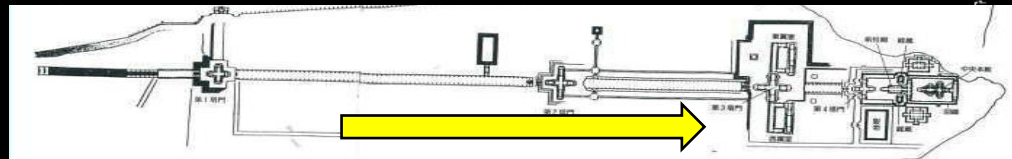
11～12世紀

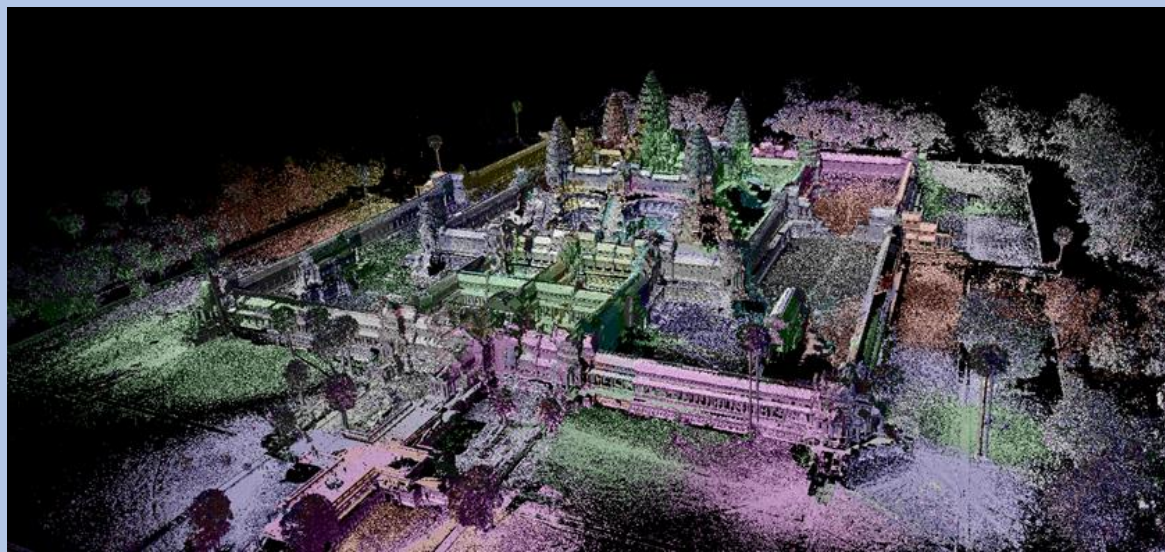
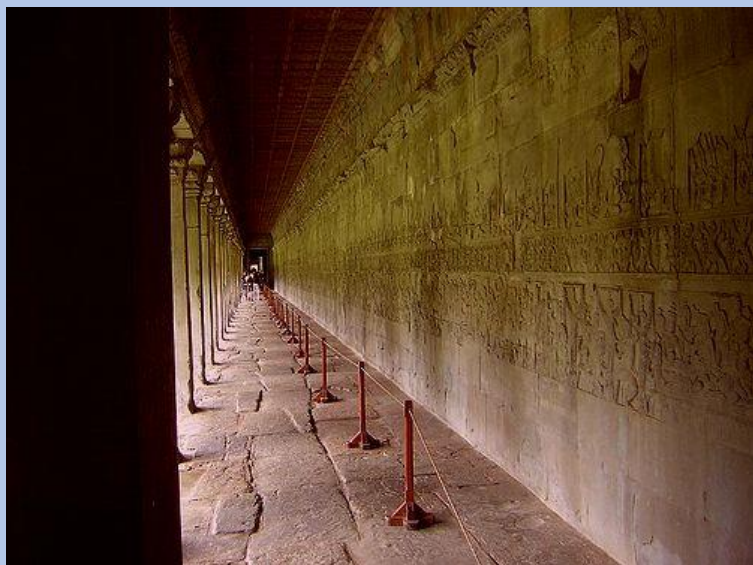
カンボジア

タイ



3次元データ → ビデオ





アンコールワット

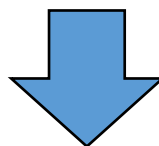
なぜワットか？



- 国のシンボル (国旗)
- 共同作業

孤立技術から共通技術へ

日本チームによるデジタル化



カンボジアチームによるデジタル化



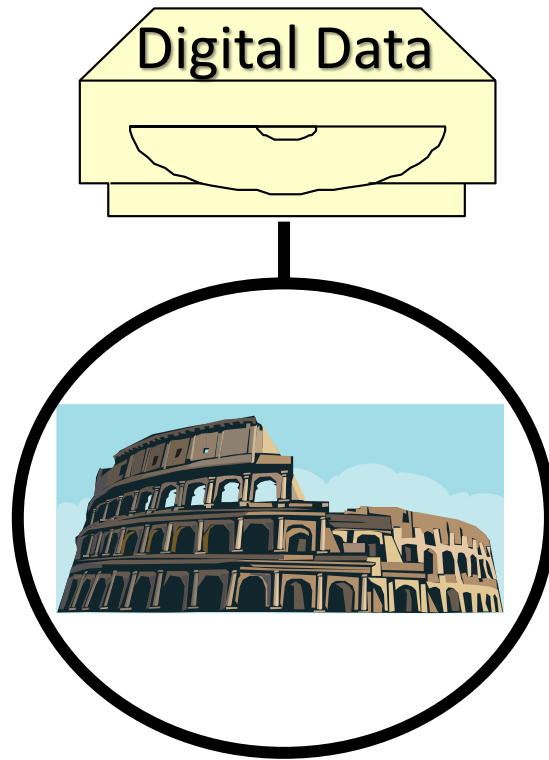
- ・モデル化は面白いし、これでVRなどが組める

⇒ 実際、凸版はこのデータを用いてバイヨンのVRシステム

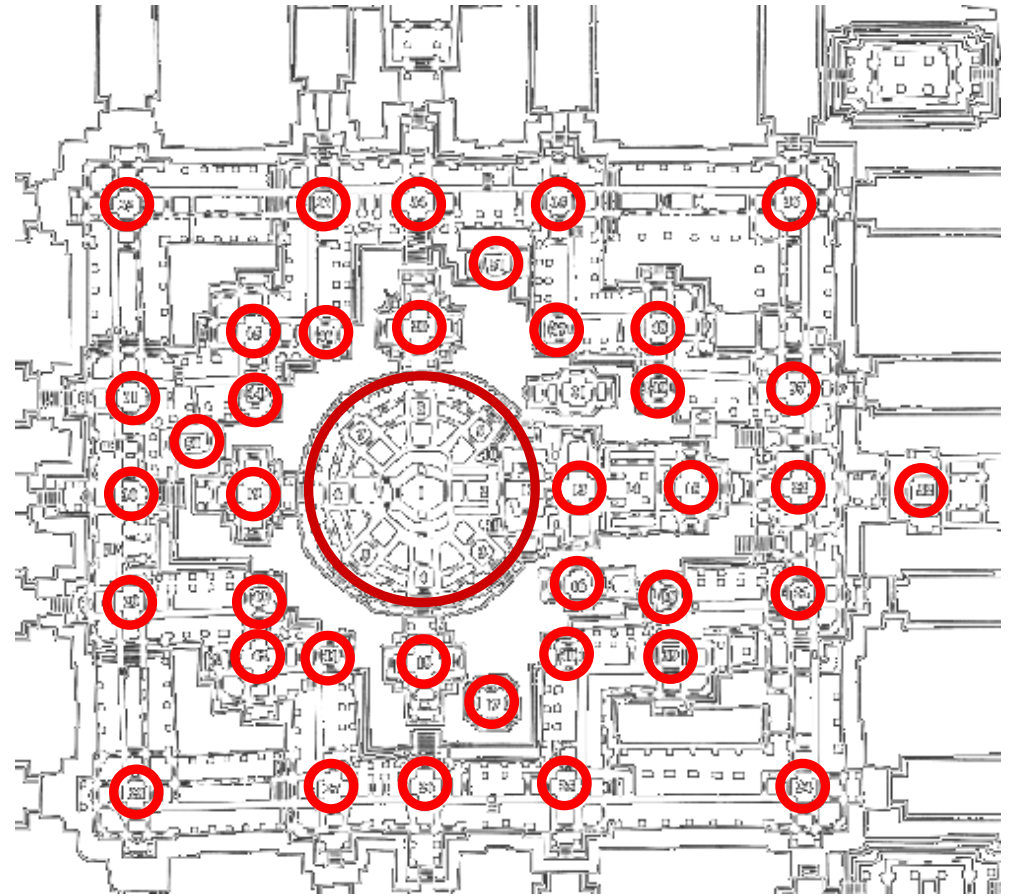
- ・そのころ、東大に情報学環が立ち上がり、生研から情報学環へ移動

⇒ 3次元データの特性を生かした文理融合の使用方法？

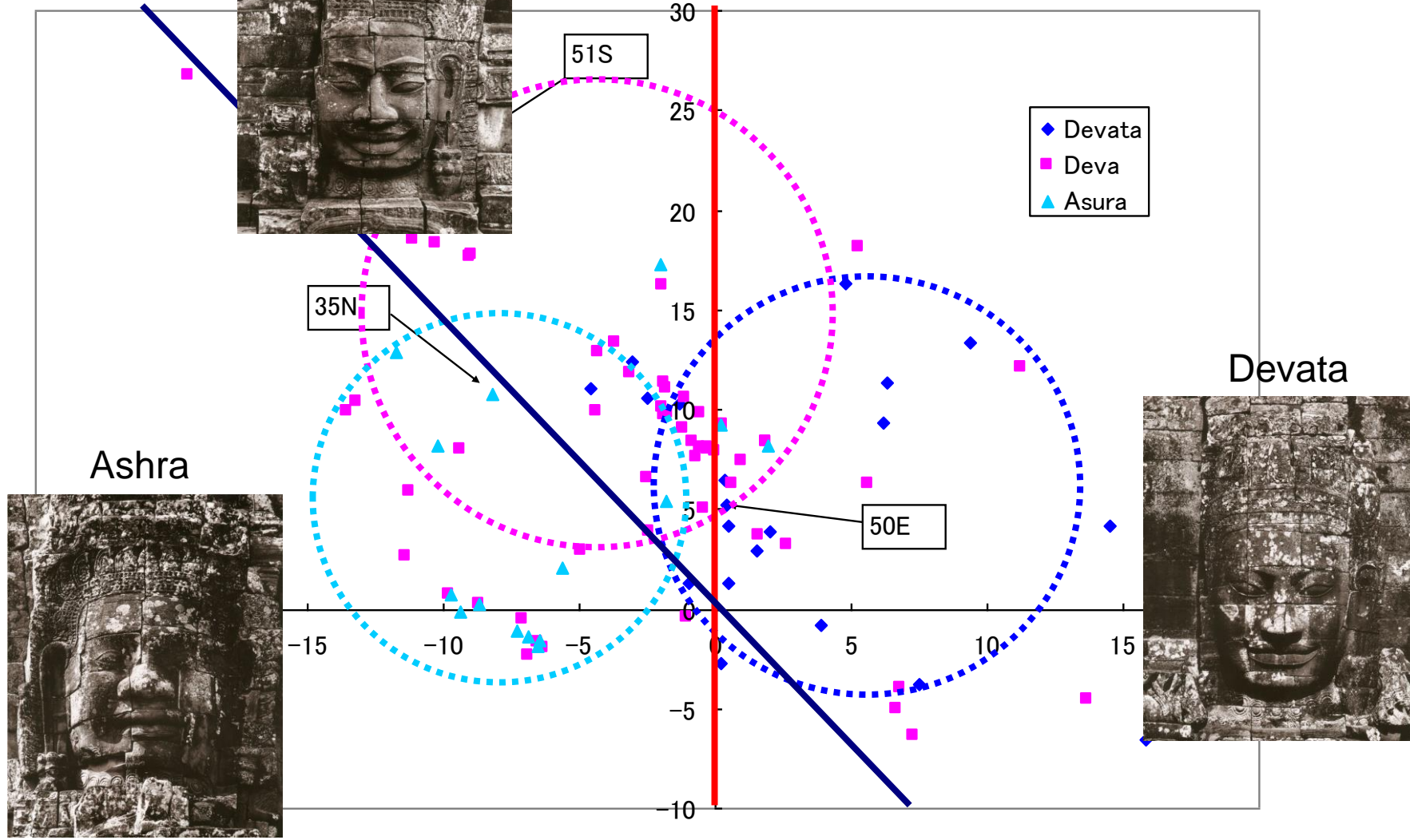
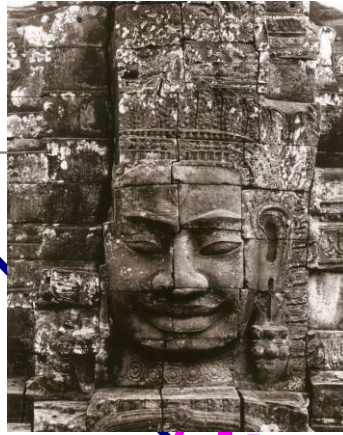
Cyber archaeology



バイヨンには、50ほどの4面塔

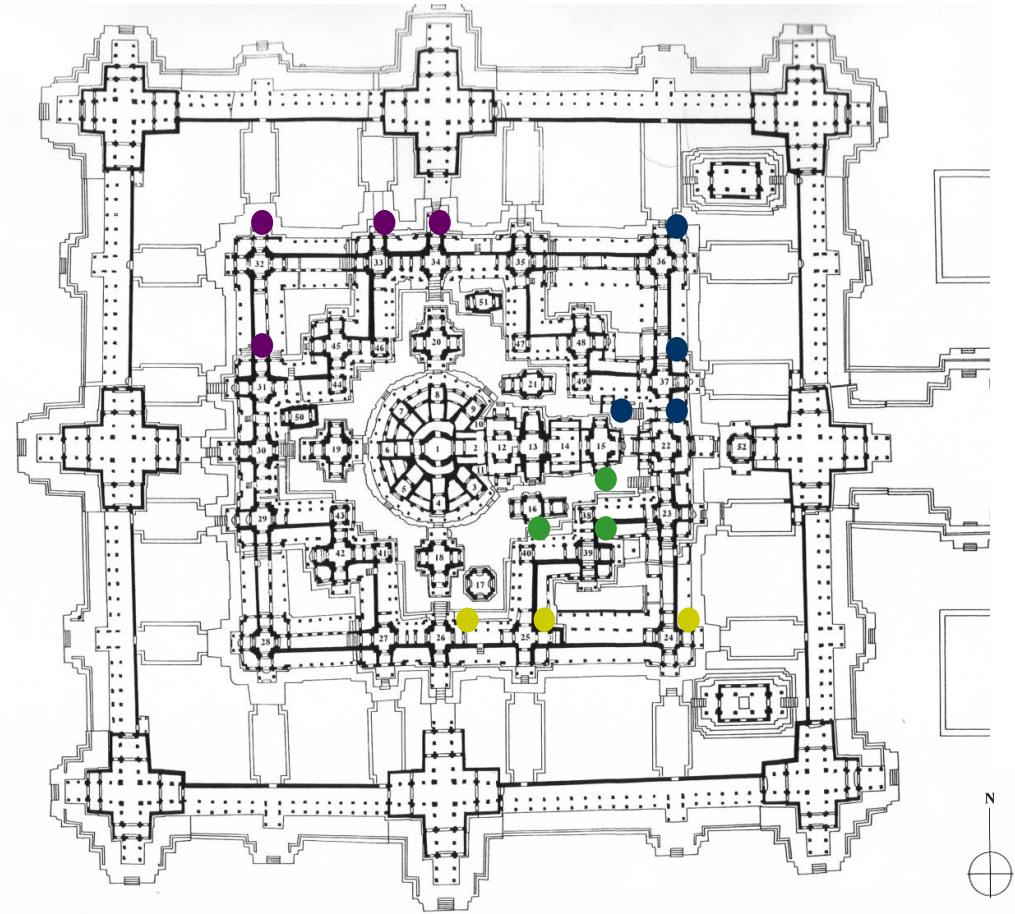


Deva



顔の類似度分布

- 似たような顔が近所にグループ
 - ⇒ 独立した石工グループが複数
子分は親分の石仏をみて削る
 - ⇒ 4つの顔類似グループ
- 4つ程度の石工グループが並行して作業

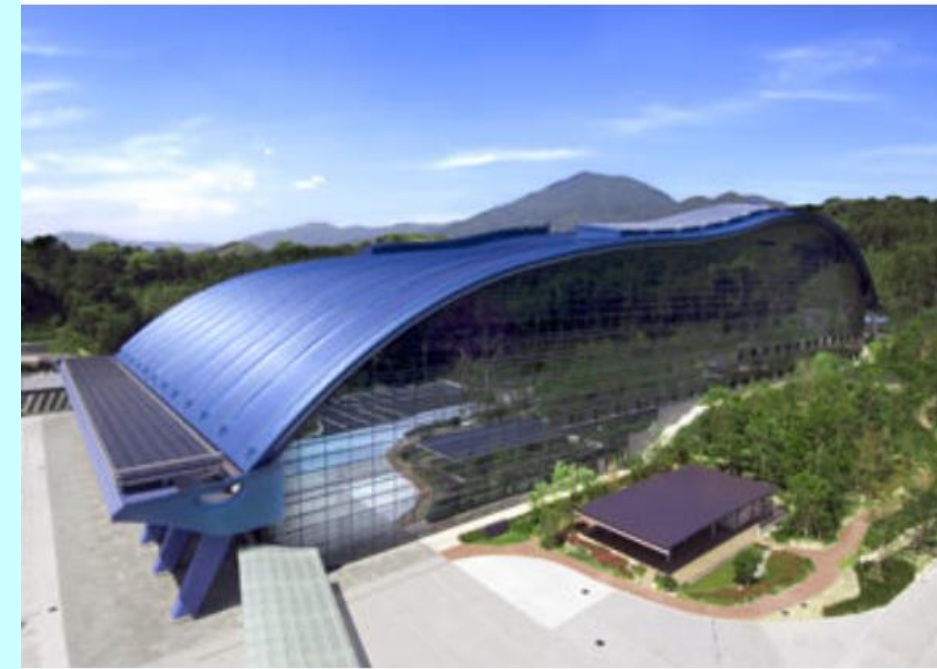
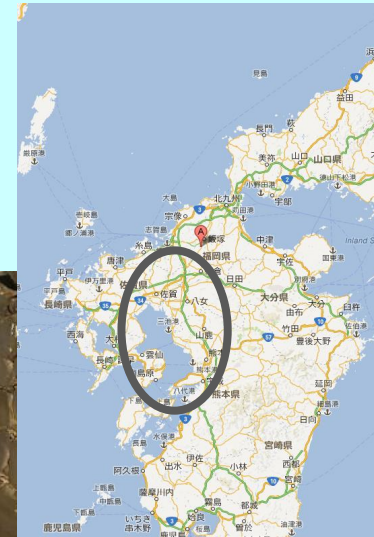


- 考古学的には、形状もさりながら、
- 光学解析の方も興味深い
- 丁度、九博の人と知り合いになり、九州装飾古墳プロジェクトを

⇒ 文部科学省直轄プロジェクト「知的資産」の開始

なぜ九州装飾古墳か？

- 筑紫平野を中心に多くの装飾古墳
- 存在があまり知られていない
- 人が入ると息(CO2)やカビで装飾色が劣化
- 閉じるとみられない(宝の持ち腐れ)
- 九州国立博物館: 常設展示物を
- 我々: 本当に色彩が重要な例





王塚 (2005)



弁慶 (2006)



日ノ岡 (2007)



千金甲 (2008)



乗場 (2009)



桜京 (2010)



田代太田 (2011)



屋形 (2012)



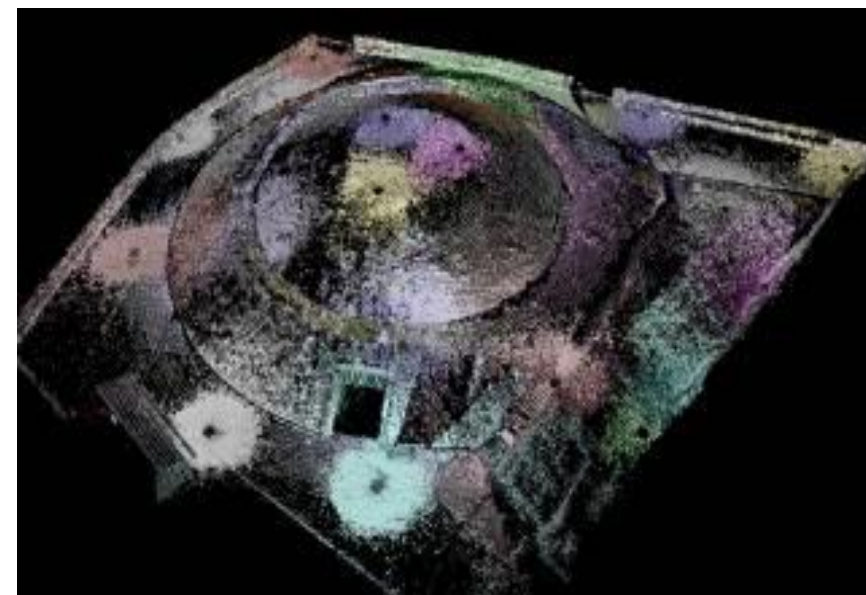
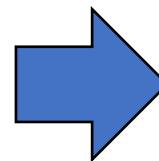
弘化谷 (2013)



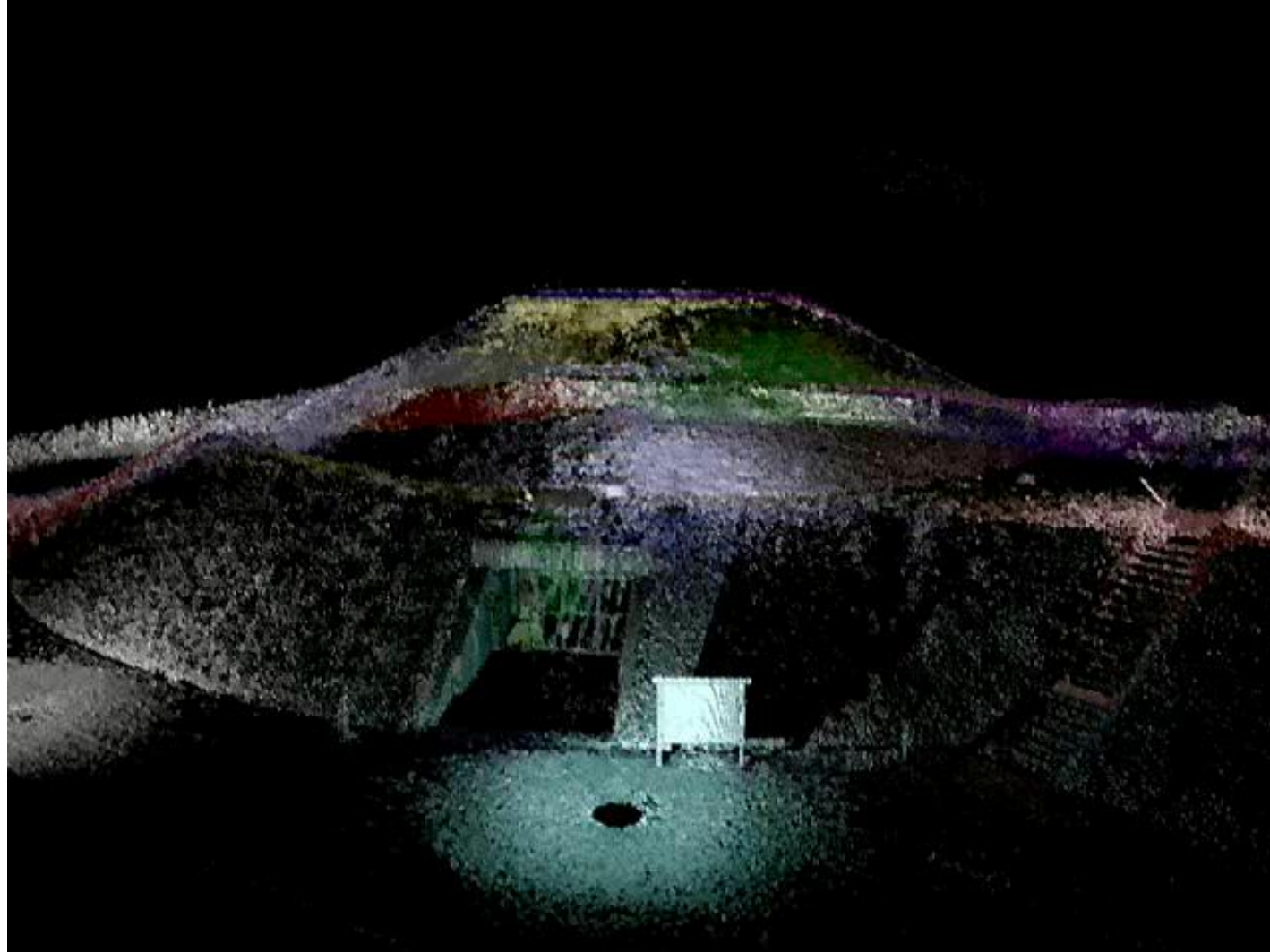
石人山 (2014)

九州装飾古墳プロジェクト

形状モデル

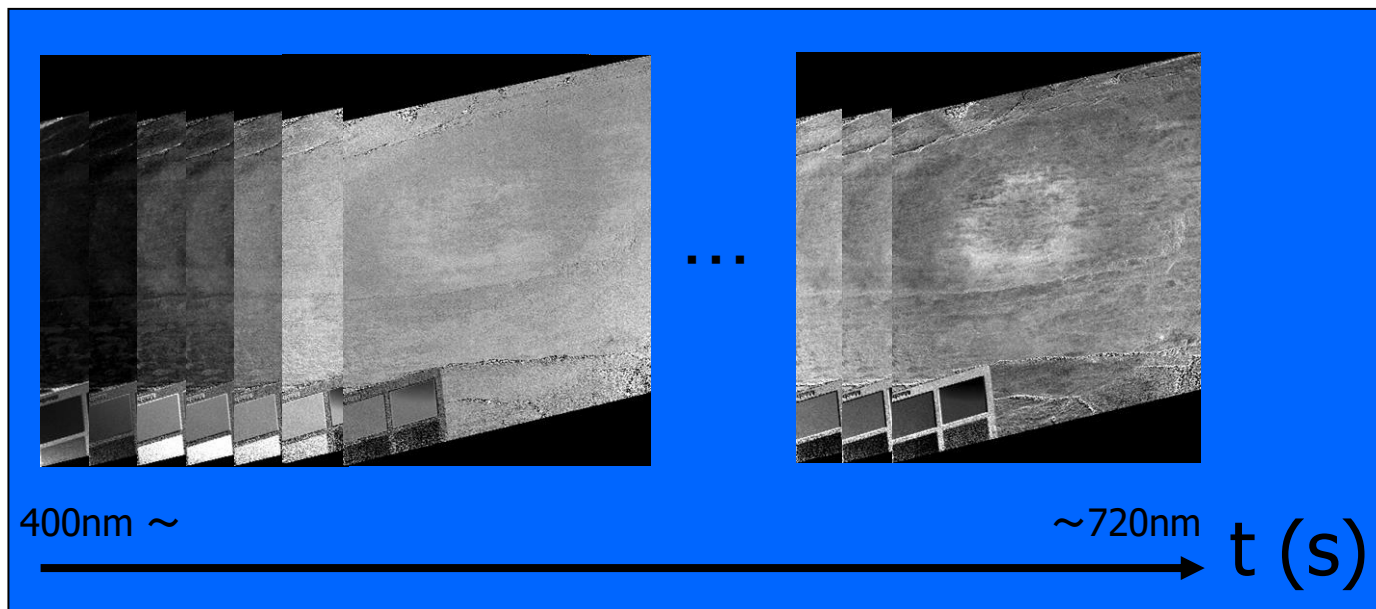
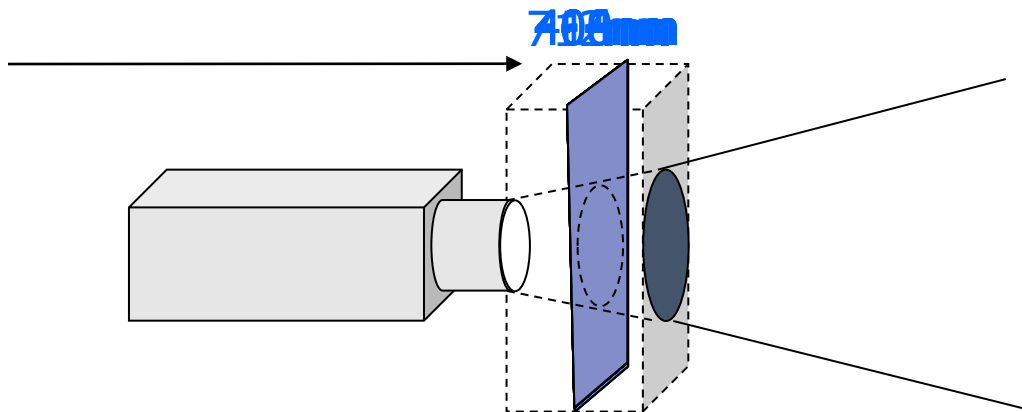


墳丘全体



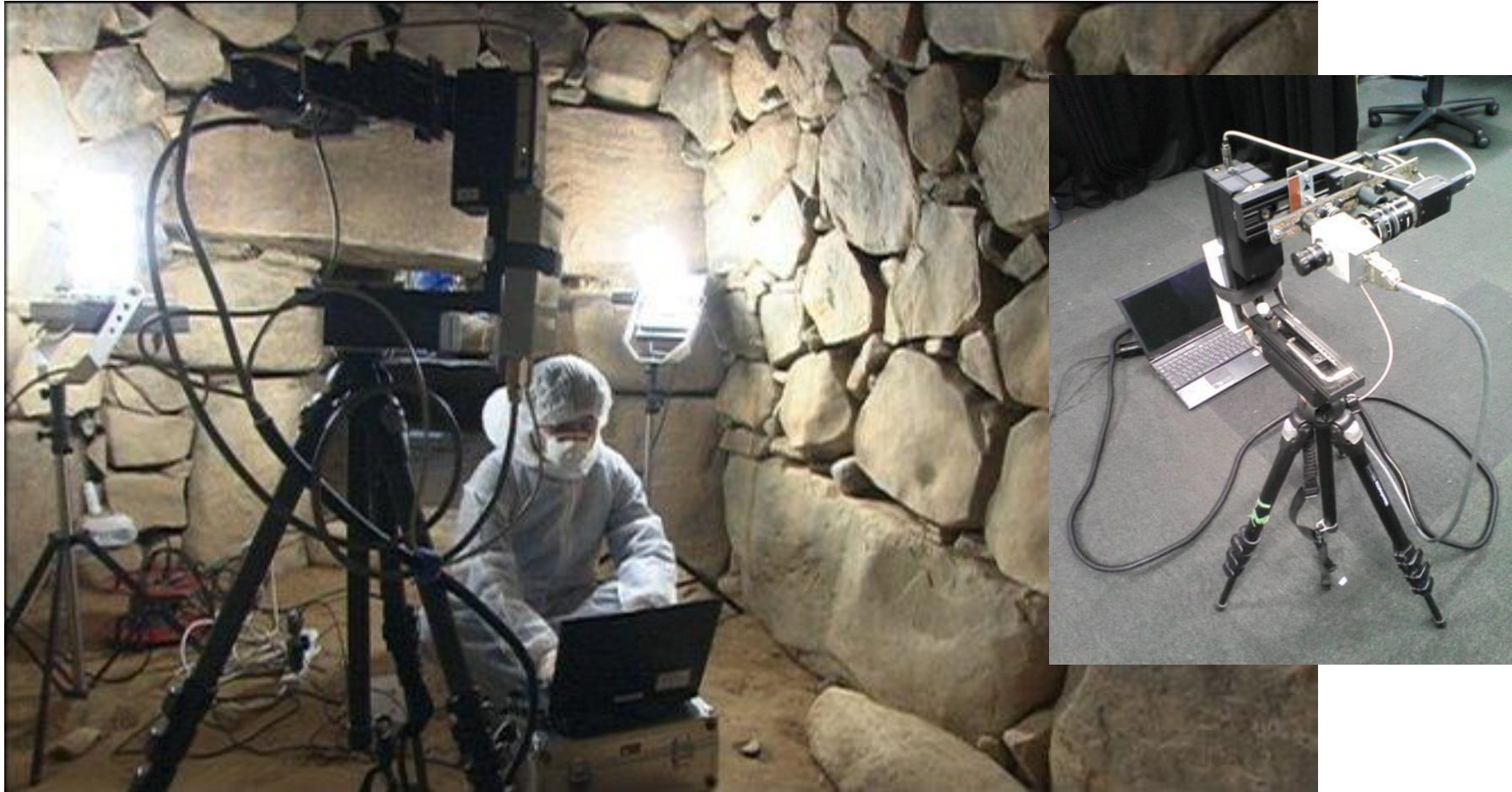
光学センサーの開発

- 電子制御により透過波長を変更できるフィルタ

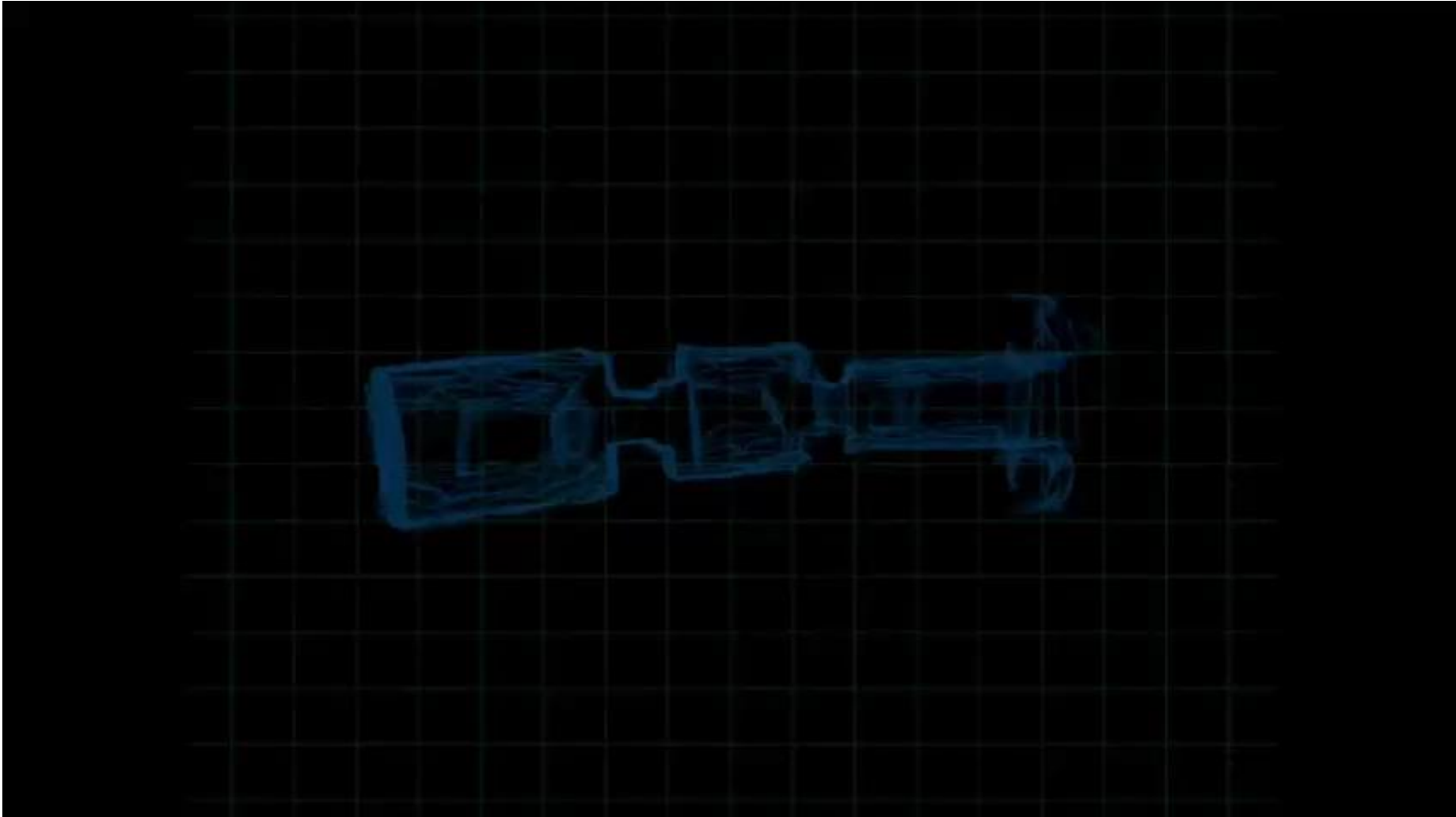


- 各画像の明るさがスペクトル強度

光学モデル



コンテンツ例





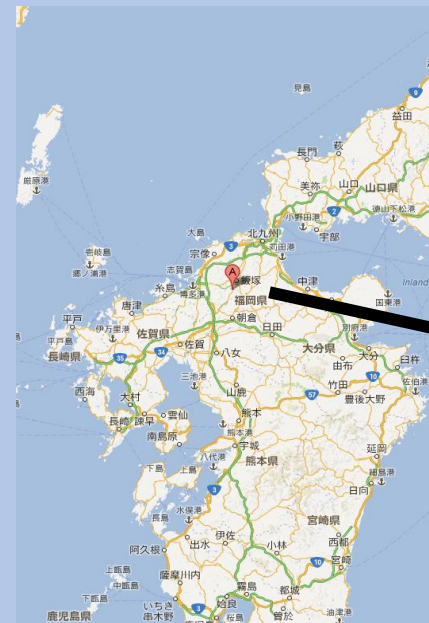
王塚古墳

6世紀

福岡県桂川町

彩色壁画（6色の顔料）

九博の常設展示コンテンツ





見えのシミュレーション

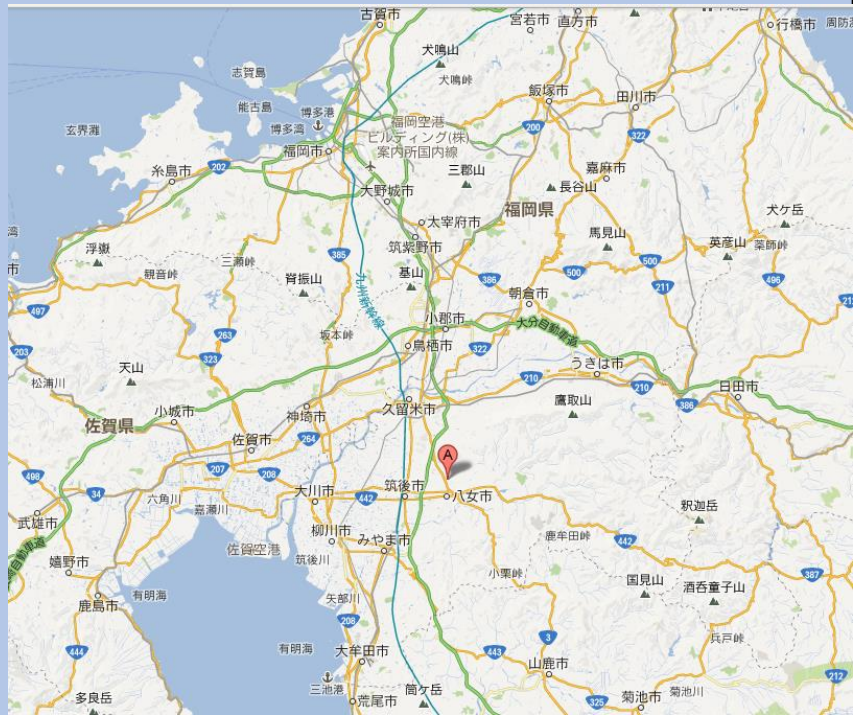
松明



太陽光



- 太陽光のもとでしか判別出来ない線
- 太陽光のもとで行われた？

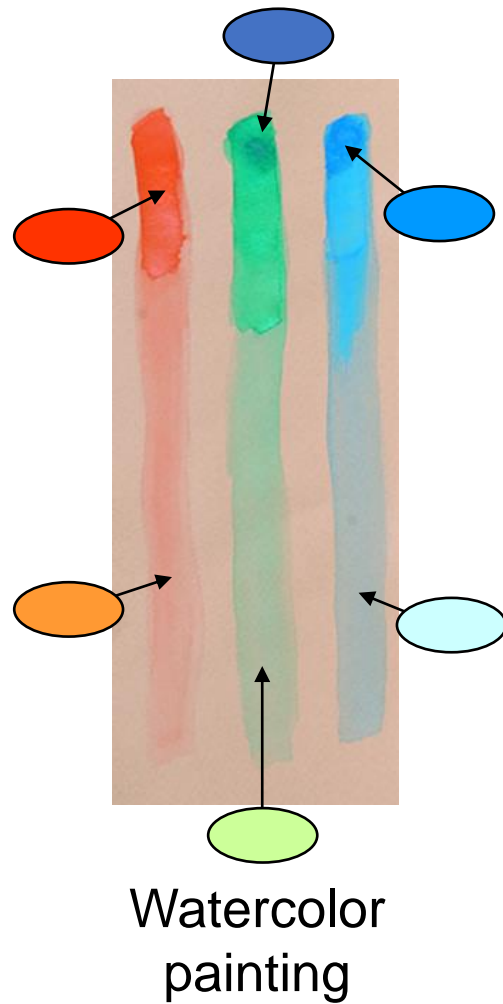


乗場古墳

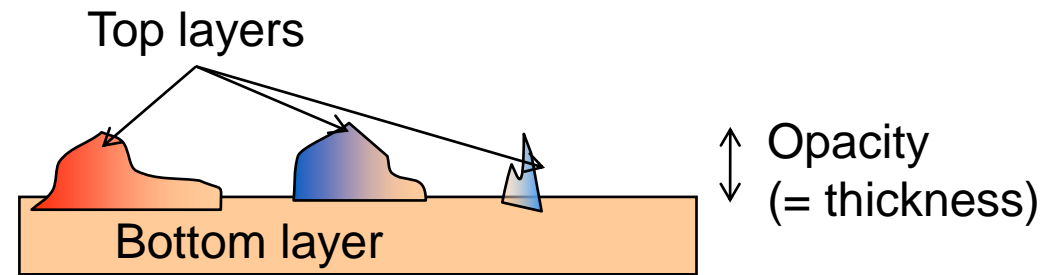
掠れた彩色 の解析



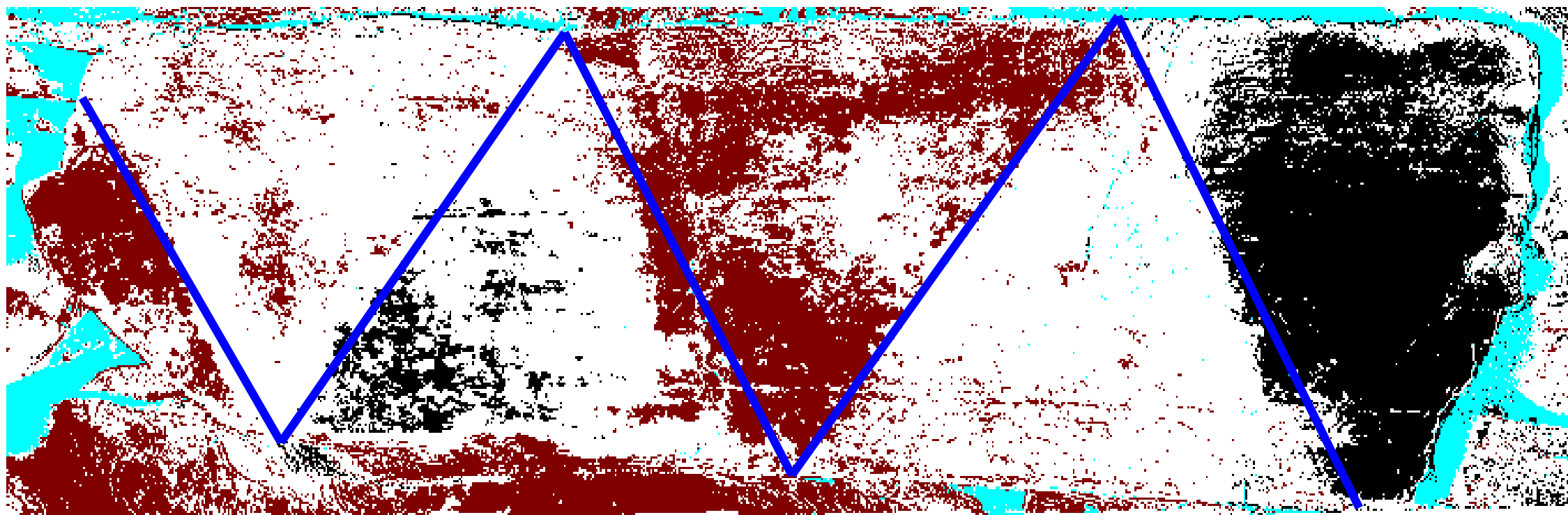
層状面のモデル



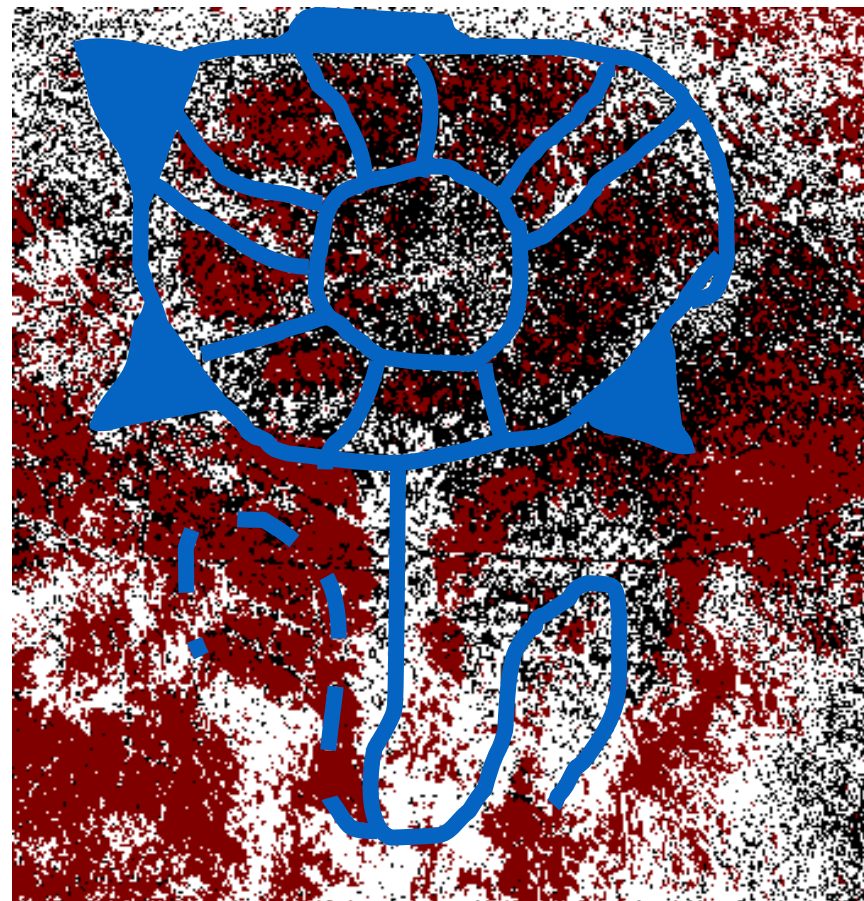
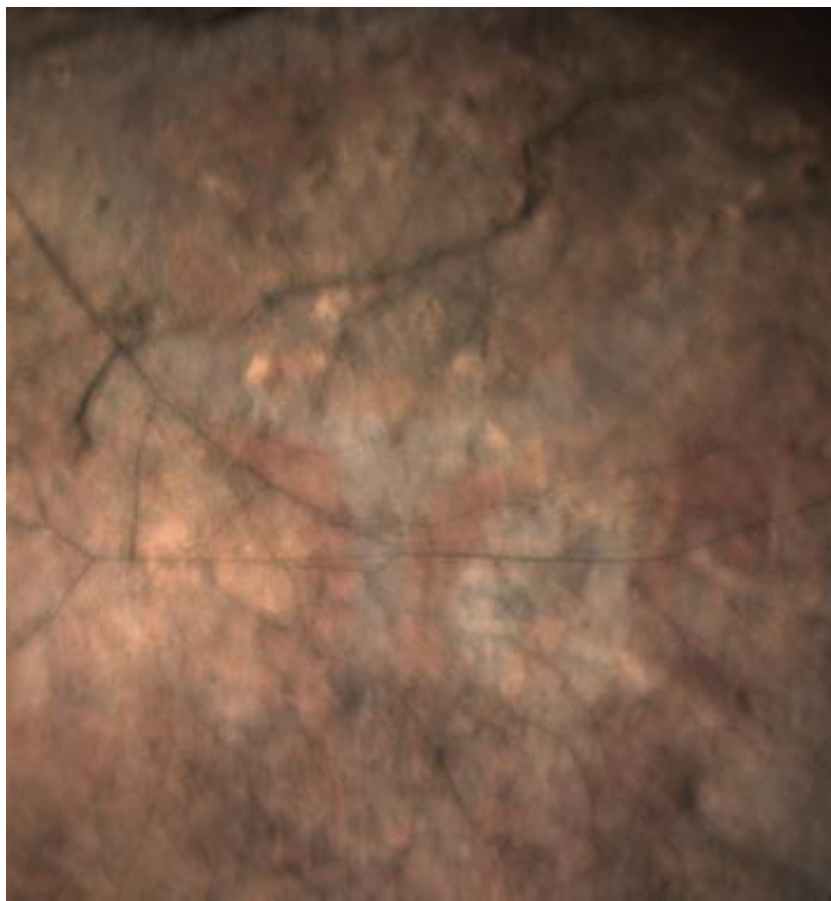
見えは、下層と上層の混合



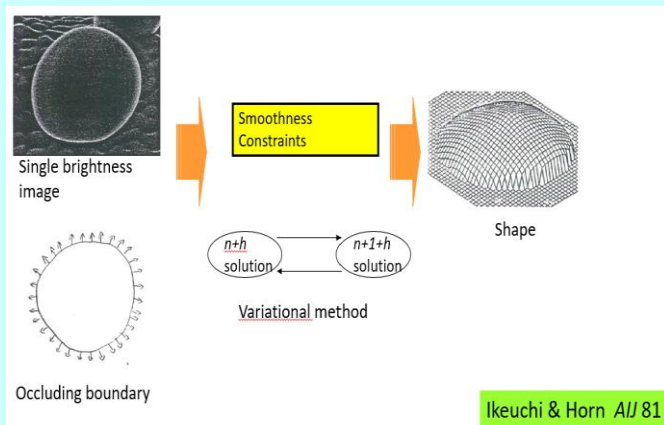
三角紋の抽出



双脚輪状文？



旅の前半は、

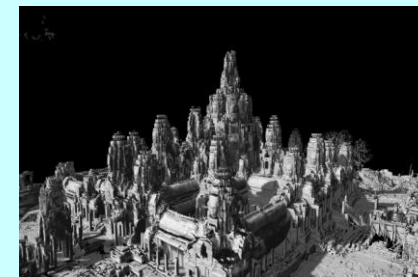


第一楽章

コンピュータビジョン

明るさ解析

E-Heritage



松明

太陽光



第二楽章

前半を終えるまえに： 開発分野

- コストを意識する開発分野
 - 日本の企業が得意
 - 若干閉塞感も
- コストを無視する開発分野
 - 「なんぼでもええから世界一をつくれ！」
 - ベンチマークシステム群
 - 工業技術のレベルを引っ張る

コストを無視する開発分野

- 軍事研究(米国・中国・韓国の大学)
 - 国家という掛け替えのないものを護る
- Xという掛け替えのないものを護る研究開発
- Xとして、
 - 人命
 - 宇宙船地球号(環境)
 - 民族のアイデンティティーたる文化

国のちから

- ハード力 = 軍事力
- ソフト力 = 文化力
- 軍事力だけの民族は、馬鹿にされる
- 文化力も以外に強力
- 軍事研究をやらないなら、せめて文化力（National Coolness）を高める研究を！

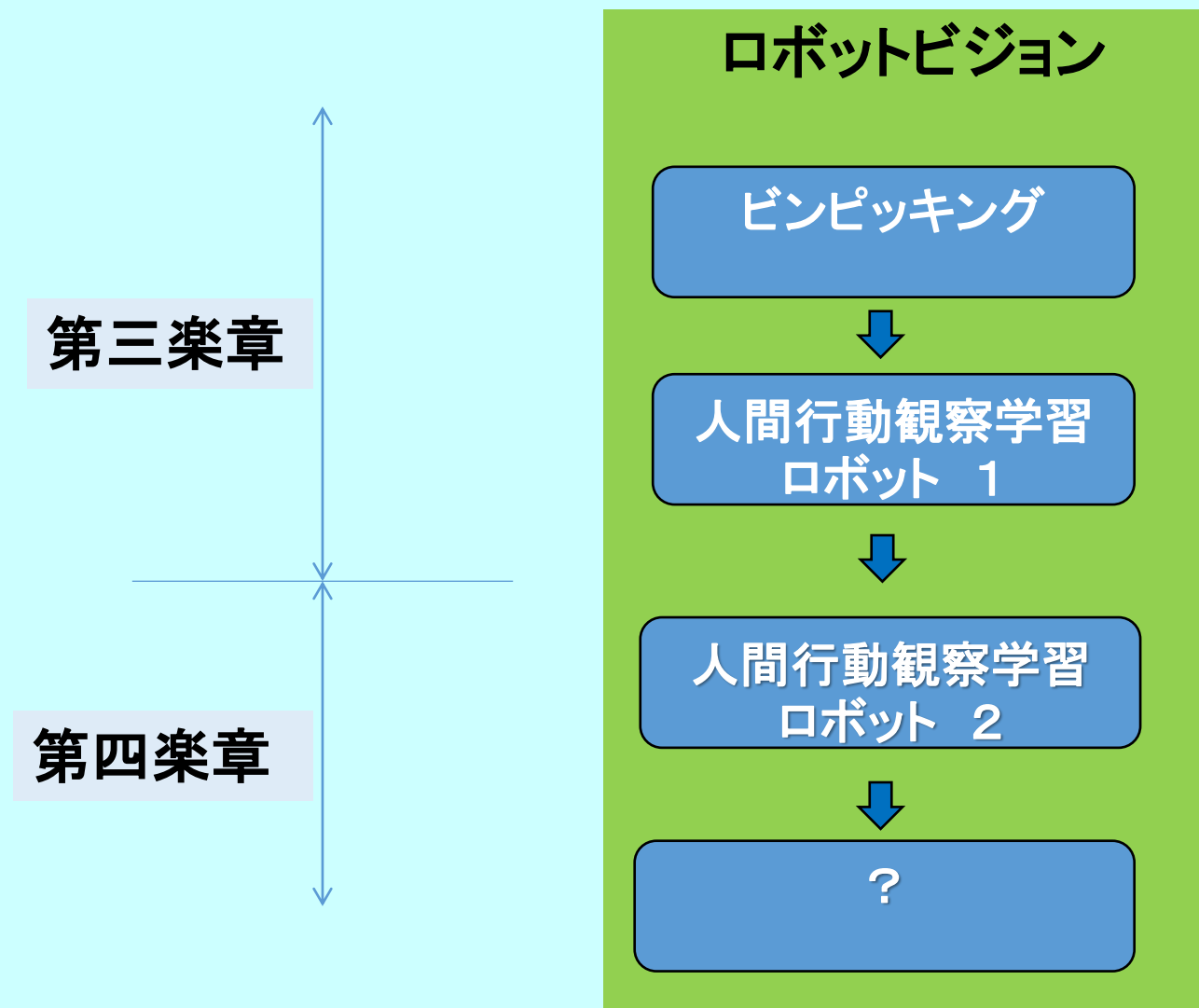
デジタル文化立国日本

- 文化をまもるための技術開発
- IT技術分野の起爆剤 ⇒ 閉塞感の打開
- 日本文化の発信 ⇒ 各国からの信頼
- 先端技術はすべて共通 ⇒ 必要ならいつでも軍事転用可能

対比

- 米国・中国・韓国
 - ・ 国家というかけがえのない物を護る技術開発
- 日本
 - ・ 文化というかけがえのない物を護る技術開発
 - ・ 文化大国
 - ・ なお且つ新しい科学技術

旅の後半はロボットビジョンでの旅



旅の出だしは、Hornからのメール

—— 1982年の春、MITより電総研に移って2年目、Horn よりメール

コピーデモから10年（ダートマスから20年）

その間の進歩を本にまとめ、その最終章をシステムで締めたい

⇒ Katsu 手伝ってくれないか？



人工知能とは？

最近の若い人

- ・ 深層学習、強化学習
- ・ ビッグデータを深層学習システムに与えて、認識システムをつくること！！

古い人

- ・ 人工知能はダートマス会議よりはじまった！！

ダートマス会議(1956)

- 主要メンバー

McCarty, Minsky, Rochester, Shannon



- 人工知能研究は、ダートマス参加者が主宰する少数の研究所にて
- MIT の Minsky は、自分の研究所の名前として “The” をつけて
The Artificial Intelligence Laboratory とした
真の人工知能研究所、唯一の人工知能研究所、これぞ人工知能研究所

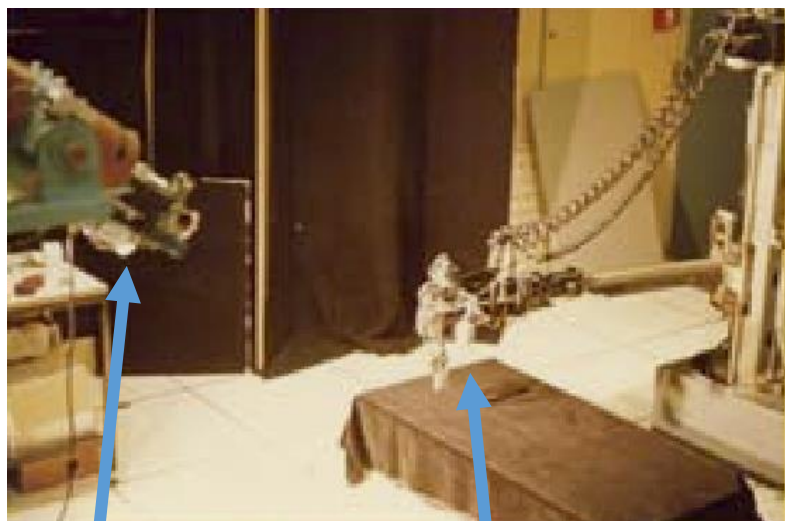
Copy demo

(真の人工知能研究所にて、ダートマス会議よりほぼ10年後)

- 目：積み木の世界を見て
- 頭：構造を理解し
- 手：そのcopyを作る



COPY



目

手



Binford (目)



Winston (頭)



Horn (手)



Minsky
(親玉)

Bin-picking demo

(真の人工知能研究所にて、ダートマス会議より、ほぼ20年)

- 目：照度差ステレオで凸凹を見る
- 頭：内部モデルと比較してドーナツツの位置姿勢を決定
- 手：把持できる方向からつかむ



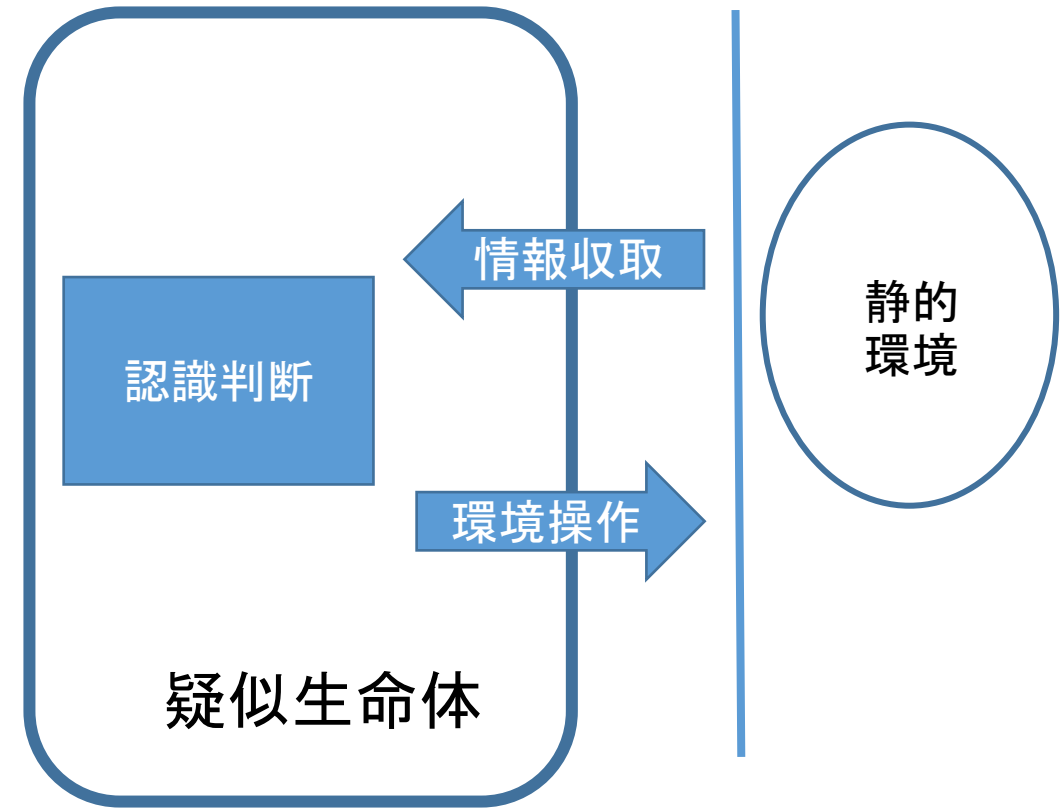
Horn & Ikeuchi *Scientific American* 84
Ikeuchi, Horn, Nagata *AI Memo* 83



長田
(金沢工大)

パイオニアたちが考えていた、人工知能とは

- Autonomous な疑似生命体
 - 目(情報収集)
 - 頭(認識判断)
 - 手・足(環境操作)
- 目的
 - 完全自律・自動
 - 無人化工場
 - 環境に人間はいない

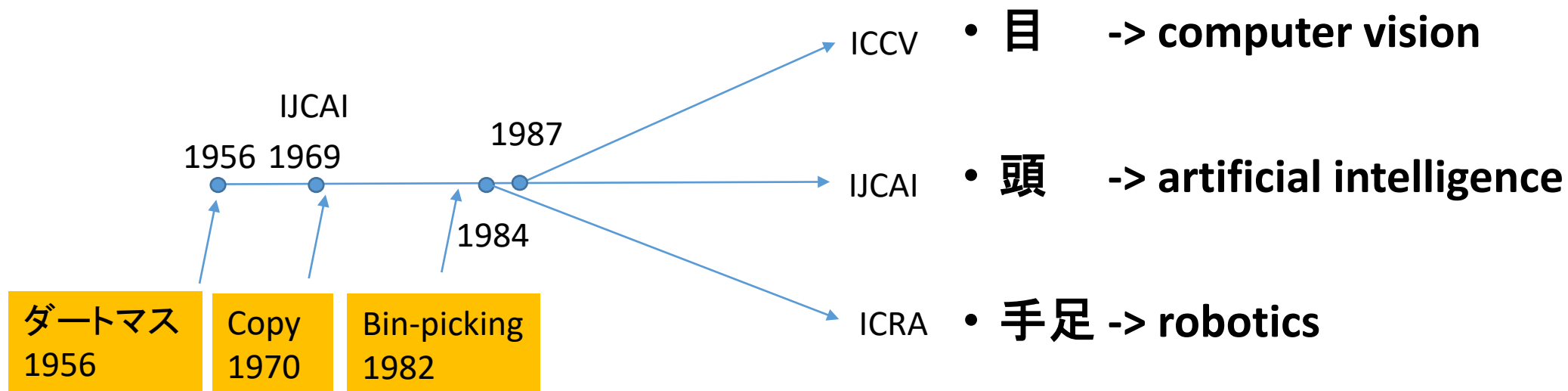


目、頭、手足の3つの要素自体大変難しい



要素還元!

- システム -> 要素
- 分割統治



Bin-picking demoの問題点

システムは知的に見えるが、
池内さんがすべての条件を考えて、条件下での行動
をプログラムに埋め込んでいる。
プログラムを通した池内さんの遠隔操縦



電総研、杉原さん

もう少し知的に、プログラムの自動生成？

⇒ 良き批判者

2つの問題（要素還元・分割統治）

- 物体認識モデルの自動生成
- 人間行動プログラムの自動生成

第3楽章 (Vivace)

人間行動観察学習ロボット

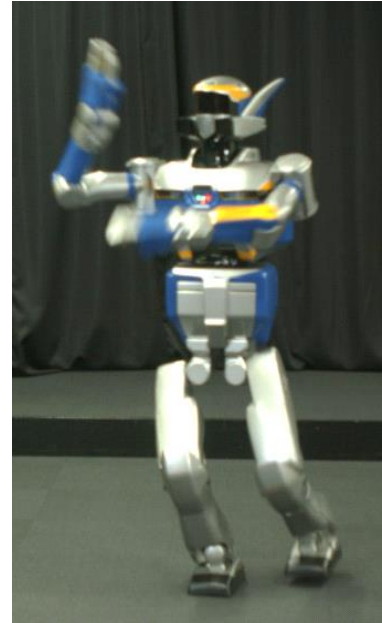
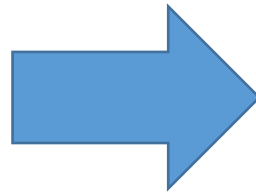
カーネギーメロン大学・ロボティクス研究所

東京大学・生産技術研究所

もしロボットが人間の行動を見て真似られれば



システムが観察

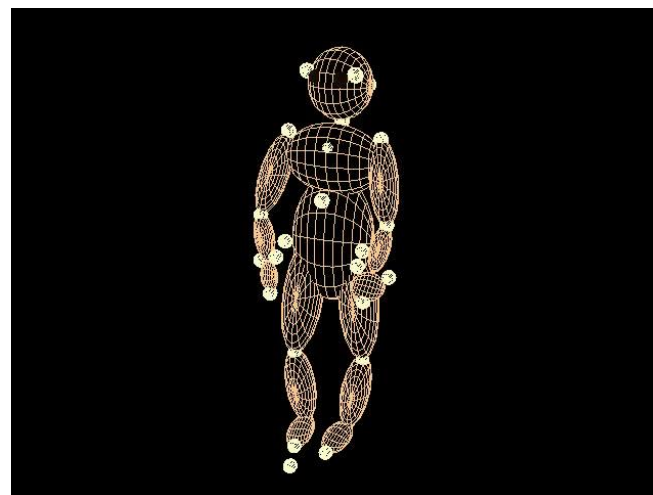
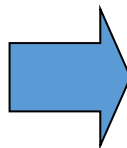


システムが動作

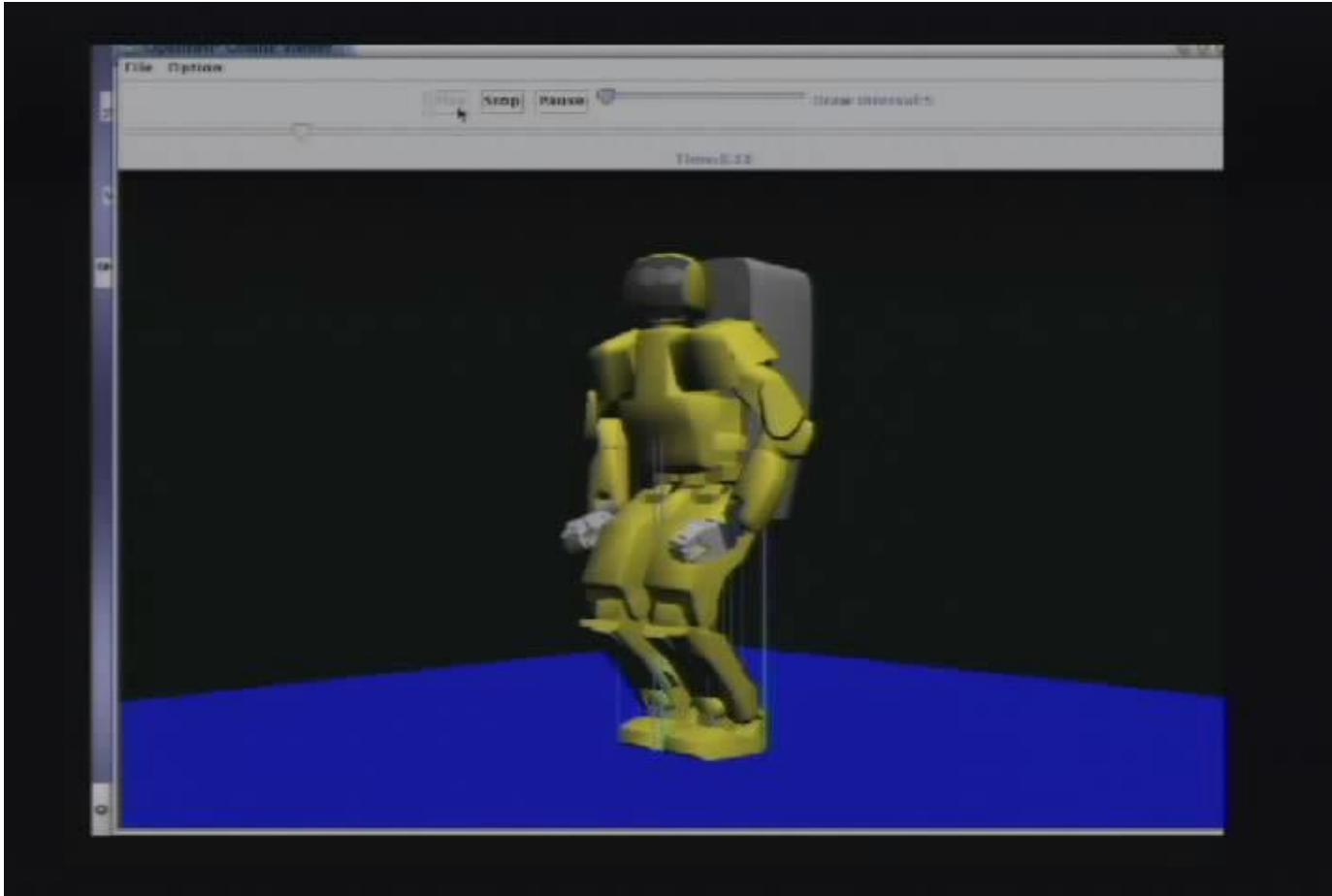
プログラムの手間がはぶける！

観察

分割統治： ビジョンの問題はさておき、
動きが完全に観察できたでしょう

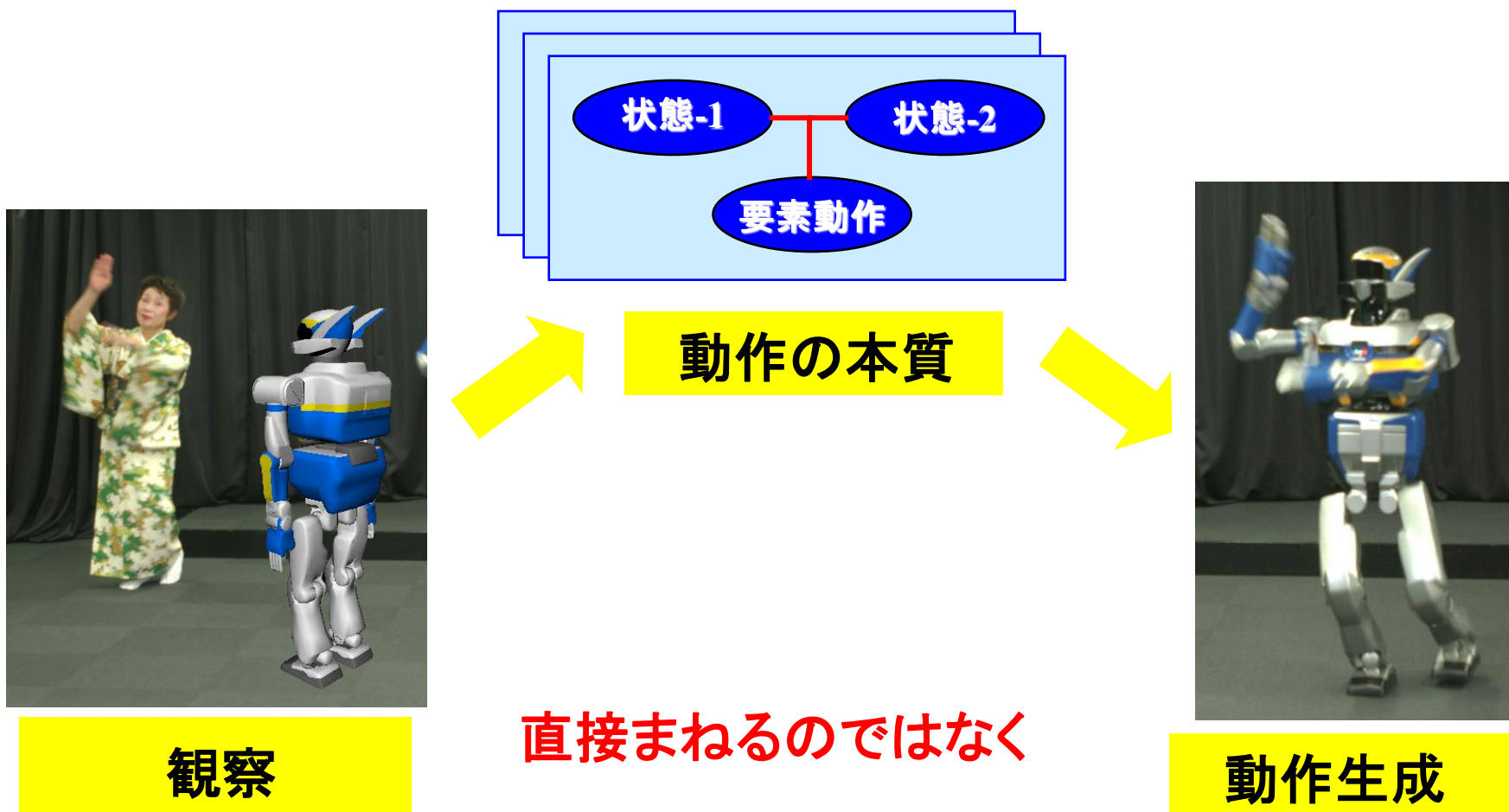


しかし、動きの本質をロボットが理解していないと、、、



産総研 動力学シミュレータ

動作の本質を知る必要



動作の本質理解は、知識のひな型を通して



- 人間は生得的に知識のひな型をもっている
- サルの赤ちゃんと人間の赤ちゃんでは目の付け所が違う
- このひな型を通して、動作を理解する

ひな形の得方

トップダウン型

- ・ 知識のひな型をはじめにロボットに用意する
- ・ ひな型はこれまでのロボット研究結果を利用

ボトムアップ型

- ・ 知識のひな型を学習によりゼロから得る
- ・ どちらかと言うと脳科学に近い

関西人的発想

- ・ 毎回ゼロからやる必要ないやん！
- ・ せっかく判ってるもんは、使わんと損やで！
- ・ トップダウンで行こ！

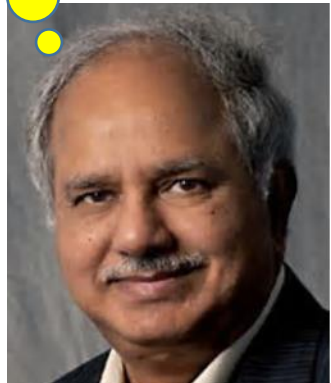
Learning-from-observation

トップダウン出発点@CMU RI



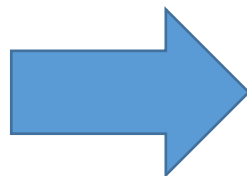
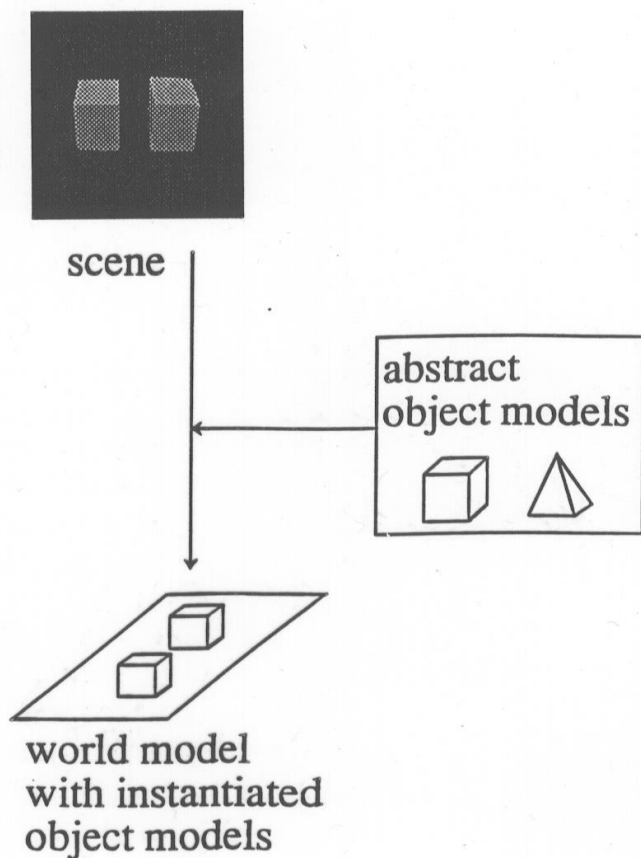
Robot World という
ロボットを寄付
してもらったけど
使わないか？

Ikeuchi & Reddy
Robotics review
1988

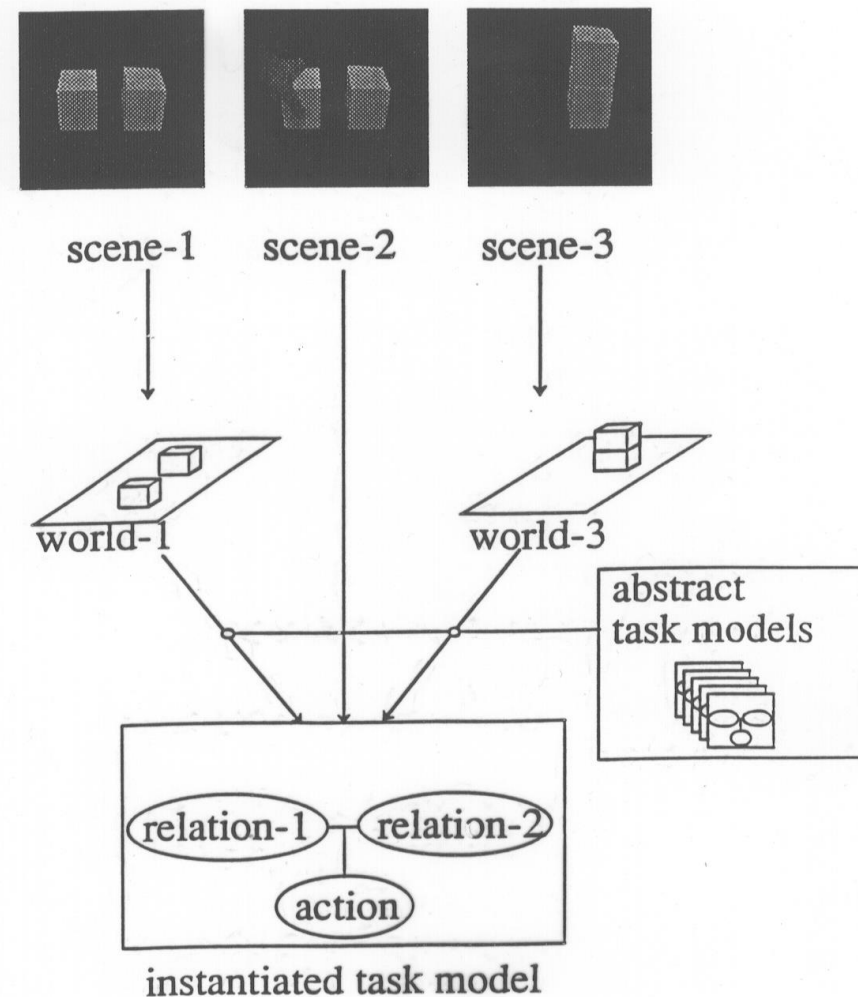


物体認識

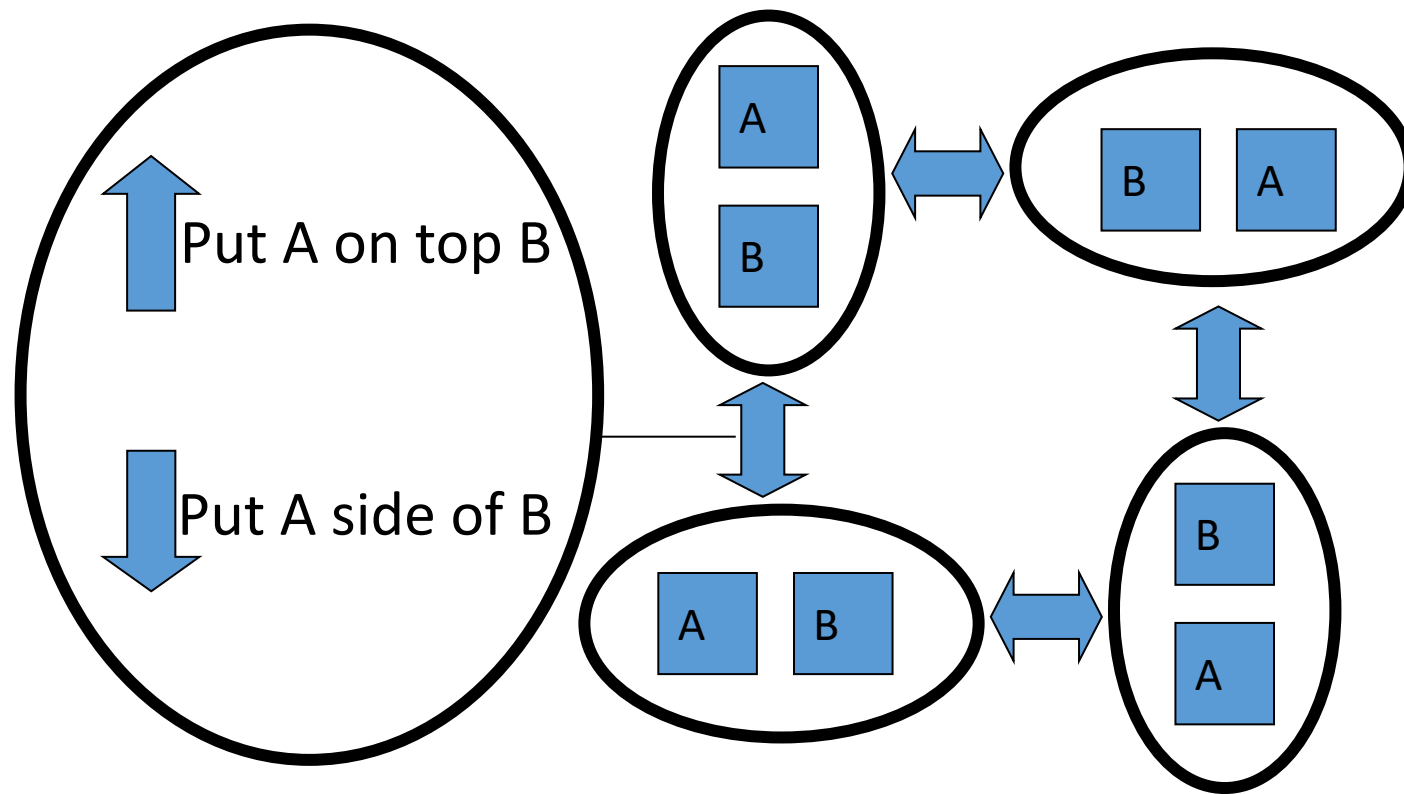
David Marr の定義



タスク認識へ



ひな型：関係・遷移・動作

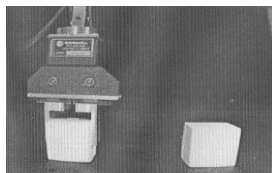


- 4状態・8遷移しかない
- 動作とは、関係(A-B)の遷移をひき起こすもの
- ひな形に、遷移に必要な動作をむすびつけ
- 物体認識の目的は、関係の遷移を見つける事

分割統治

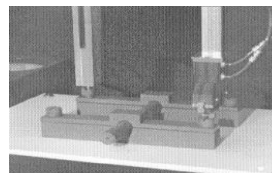


2つのブロック

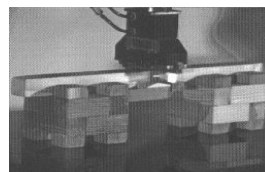


1988

機械部品



1990



多面体

2000



柔軟物(ひも)

ダンス



2010

博士論文11本:

小川原(和歌山大)、工藤(電通大)、佐藤啓(京都先端大)、中岡(産総研)

Kang (Microsoft)、Paul(スタートアップ)、高松(Microsoft)、白鳥(Meta)

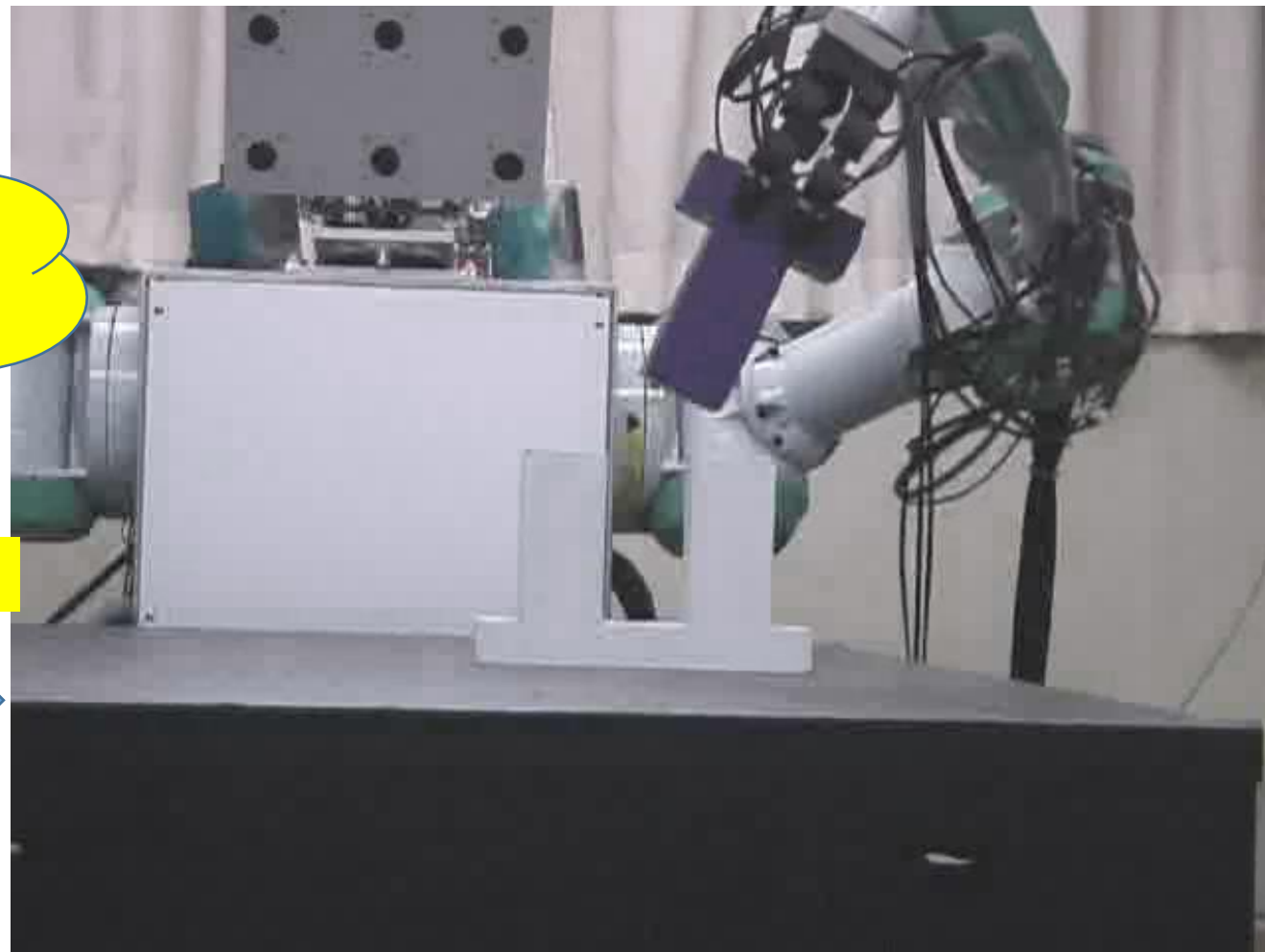
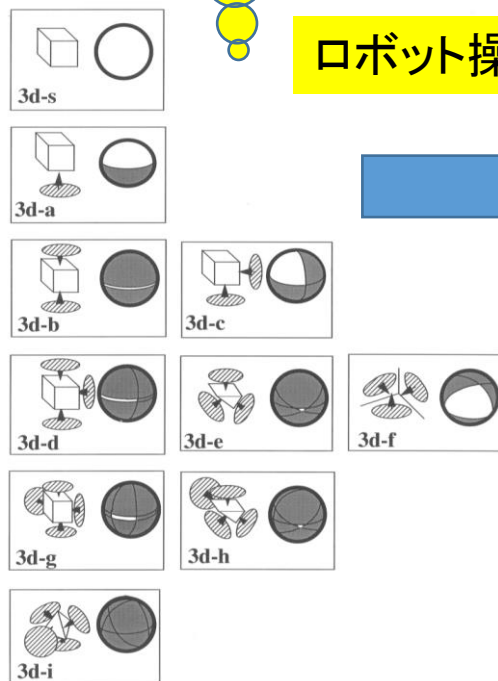
岡元(ファナック)、ミルク(スタートアップ)、マノジ(スタートアップ)

(+ 末広、川出、三浦)

多面体

Kuhn Tucker Theory
多面体接触関係

ロボット操作



Ikeuchi &
Suehiro
TRO94



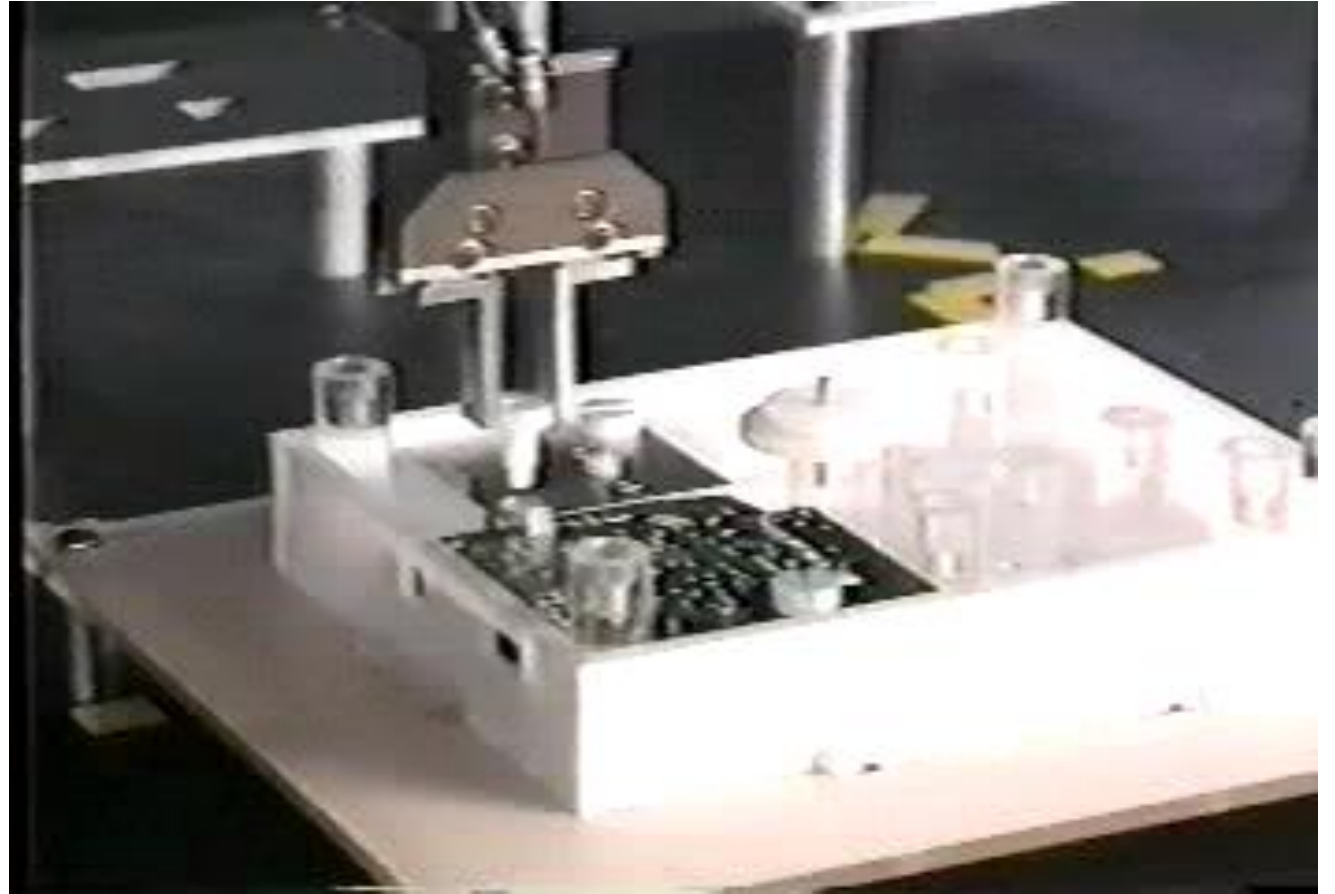
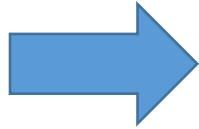
Takamatsu
et. al. *IJRR*
2002



機械部品でもお なようなひな型

機械部品間の
はめ合い関係

ロボット操作

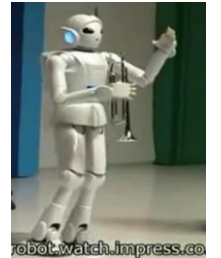


Miura and Ikeuchi,
IEEE-PAMI 1998

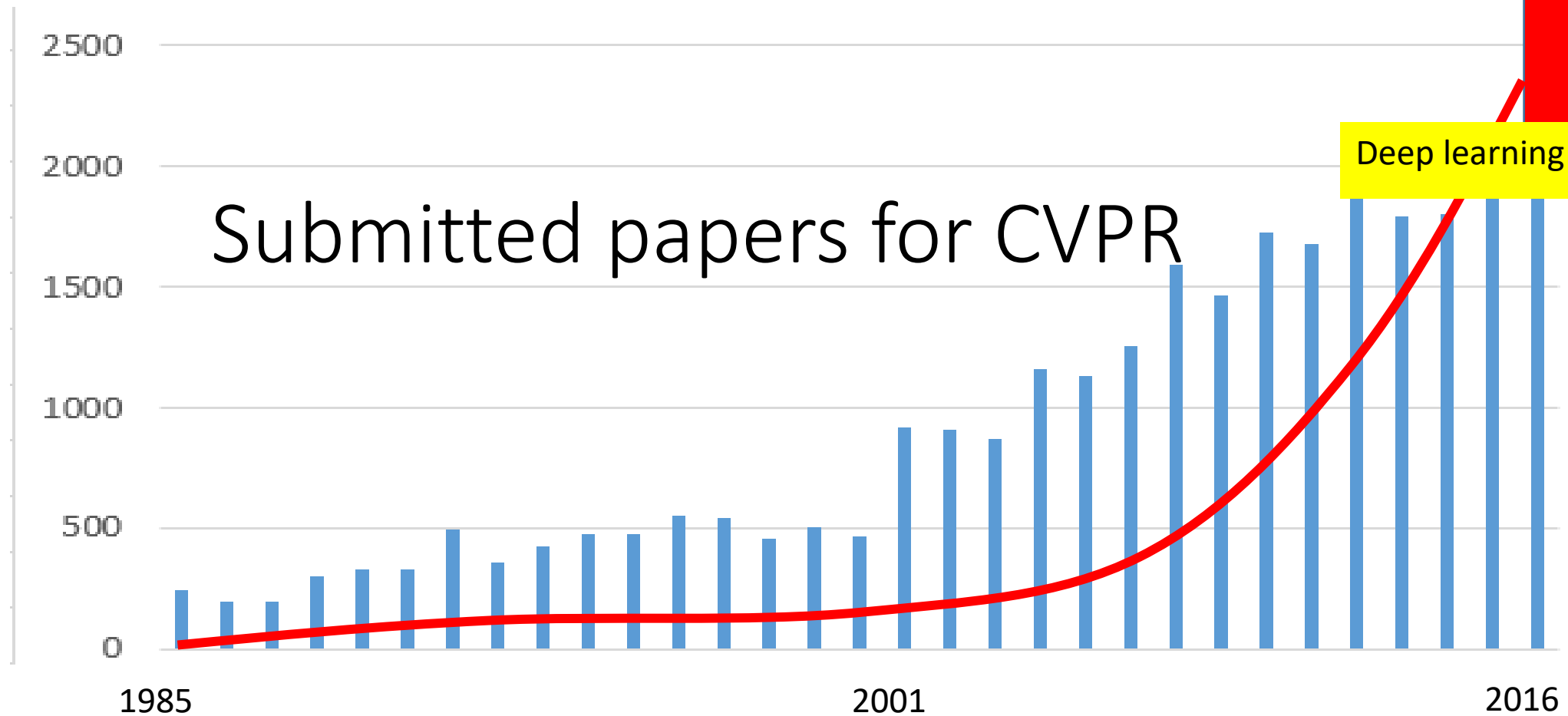


そうこうしている間に

ロボティクスにおけるカンブリア爆発 (ダートマスより 60年後)



コンピュータビジョン(CVPR)と機械学習(NIPS) におけるカンブリア爆発



Connectionism

カンブリアン爆発と大淘汰

- カンブリアン爆発
 - 生物体: 視覚器官の獲得
 - CV: 深層学習の獲得
- 大淘汰
 - 大半が淘汰

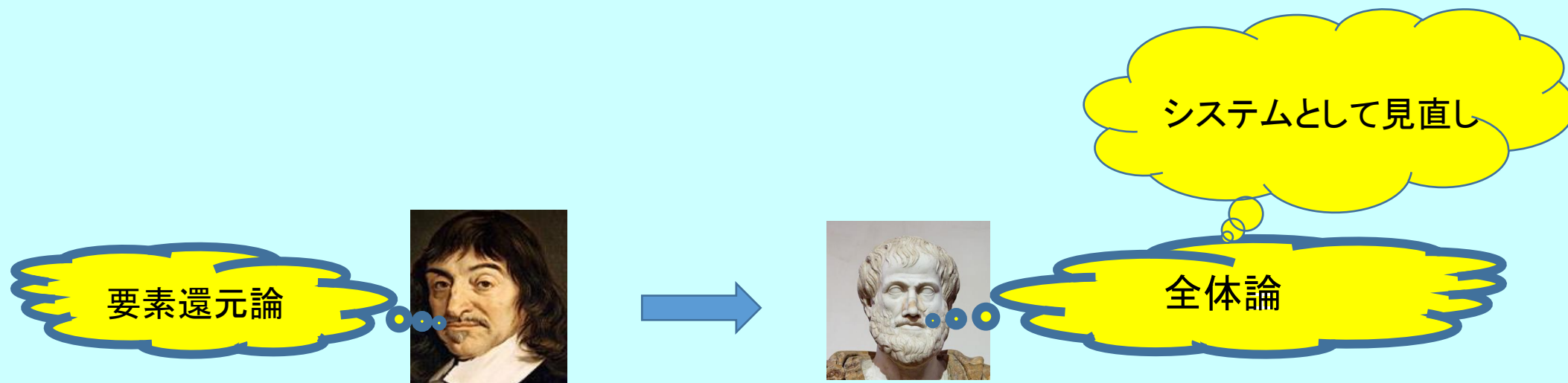
第4楽章 (Adagio)

パラダイムシフトとロボット魂

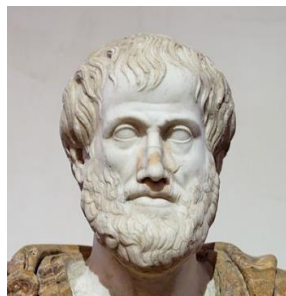
米国マイクロソフト・研究開発本部

カンブリア爆発を乗り越えるために パラダイムシフトが必要

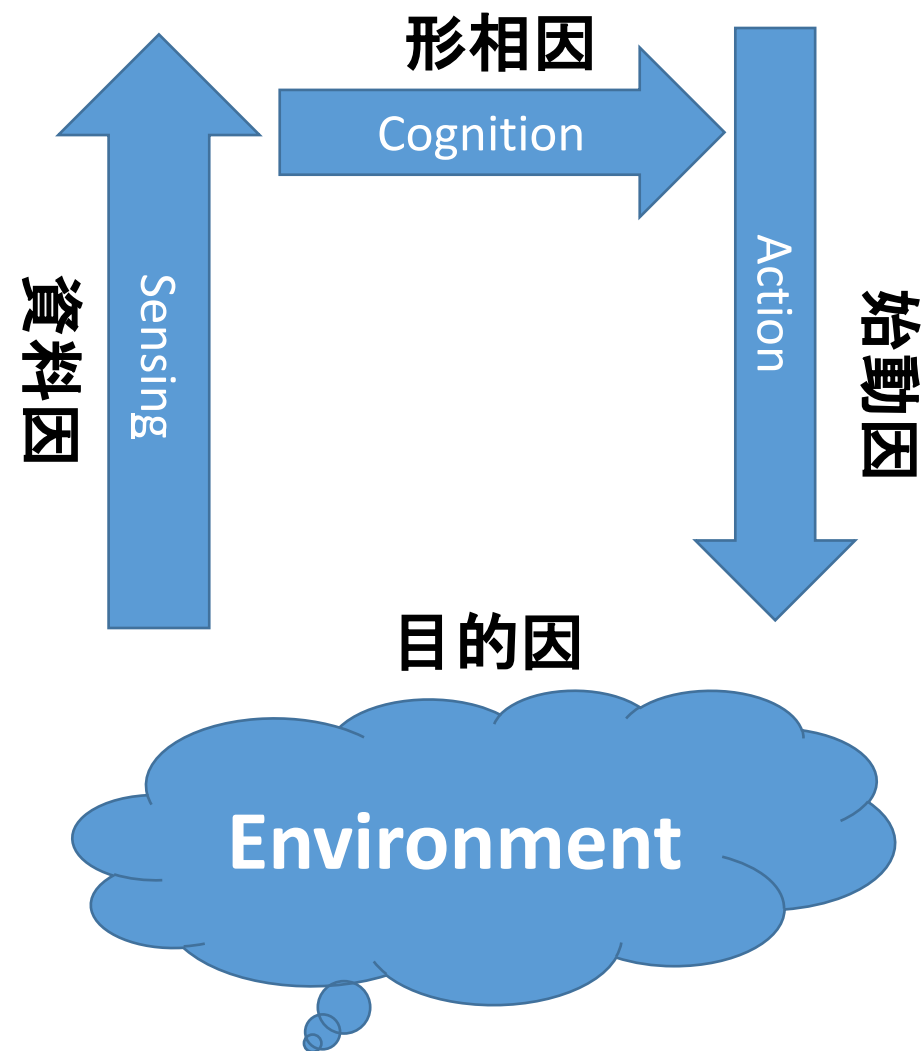
- 20世紀は、要素還元論の時代
 学問が細分化され過ぎたのではないか
- 全体論にもとづきシステムとして見直すべきではないか

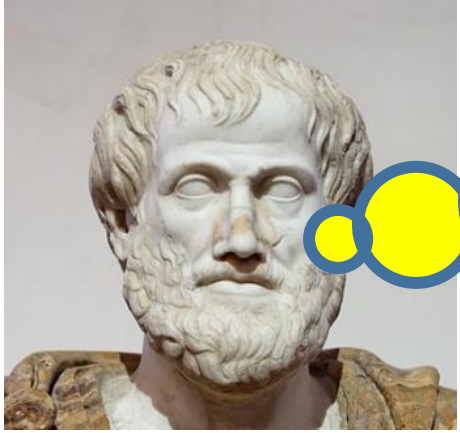


四原因論



- 質料因: 素材は何からできているか
- 形相因: それは何であるか
- 始動因: 動きをあたえるのは何か
- 目的因: 何のために成立するのか
- 目的因、環境とのインターラクションがミッシングリンク





- ・システムの生き残りのためには、環境への順化
- ・まず、環境を定義
- ・その環境に対して、システムの最適化

-> 目的因!

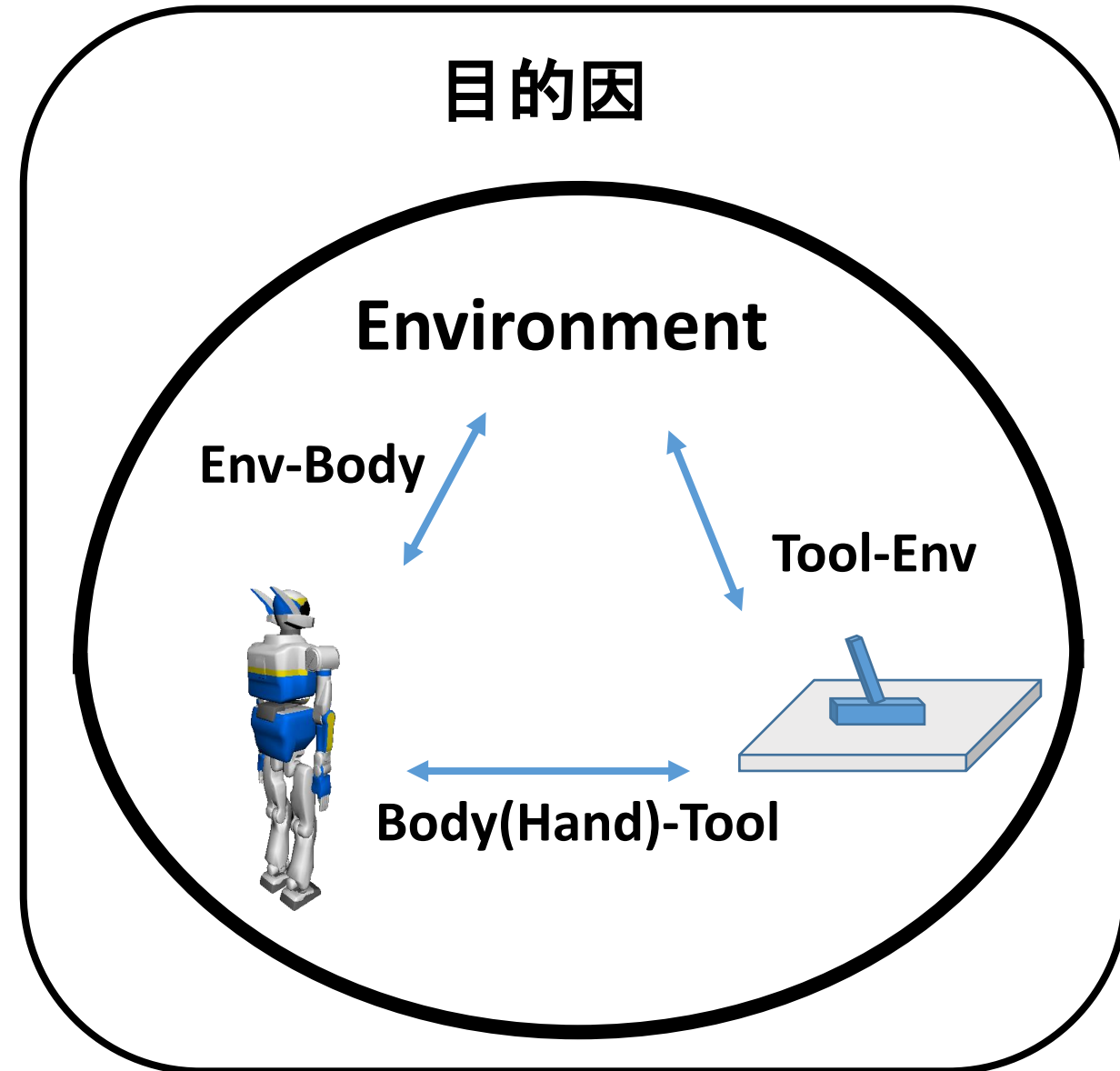
人間環境

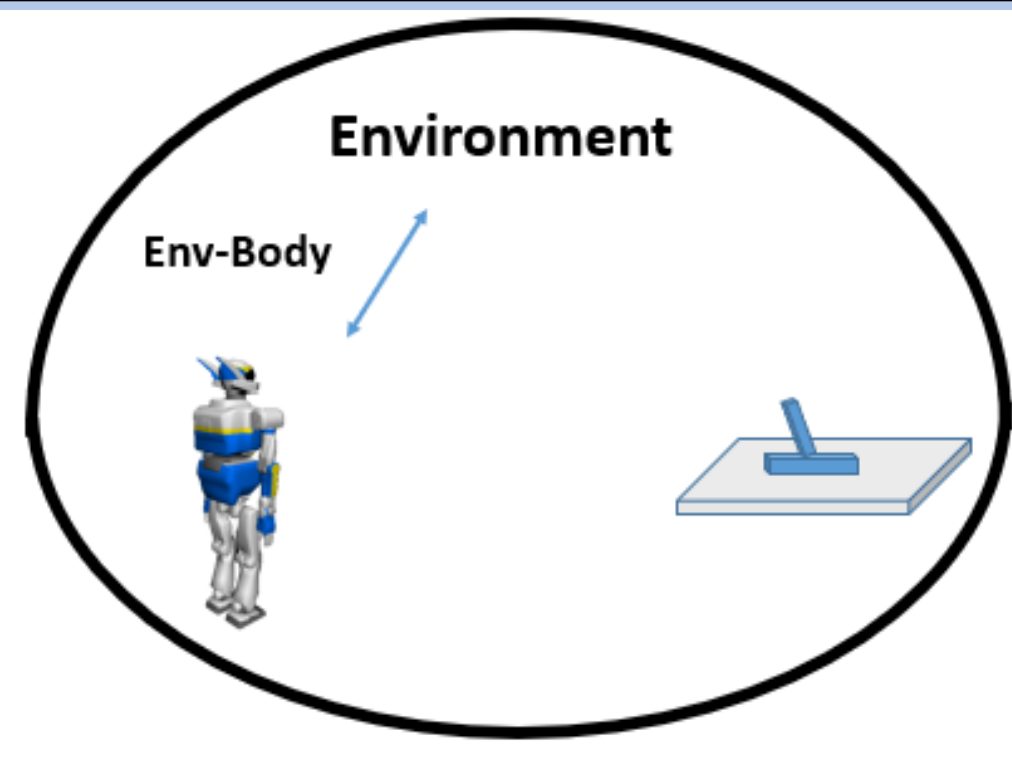
- 乱雑
- 周囲に人間
- 周囲の人との常識の共有



目的因を通した常識

- 環境と体の間の常識
 - その目的のためこういった姿勢をとるのか
- 体と道具の間の常識
 - その目的のためこういった把持をとるのか
- 道具と環境の間の常識
 - その目的のためこういった動きをするべきか

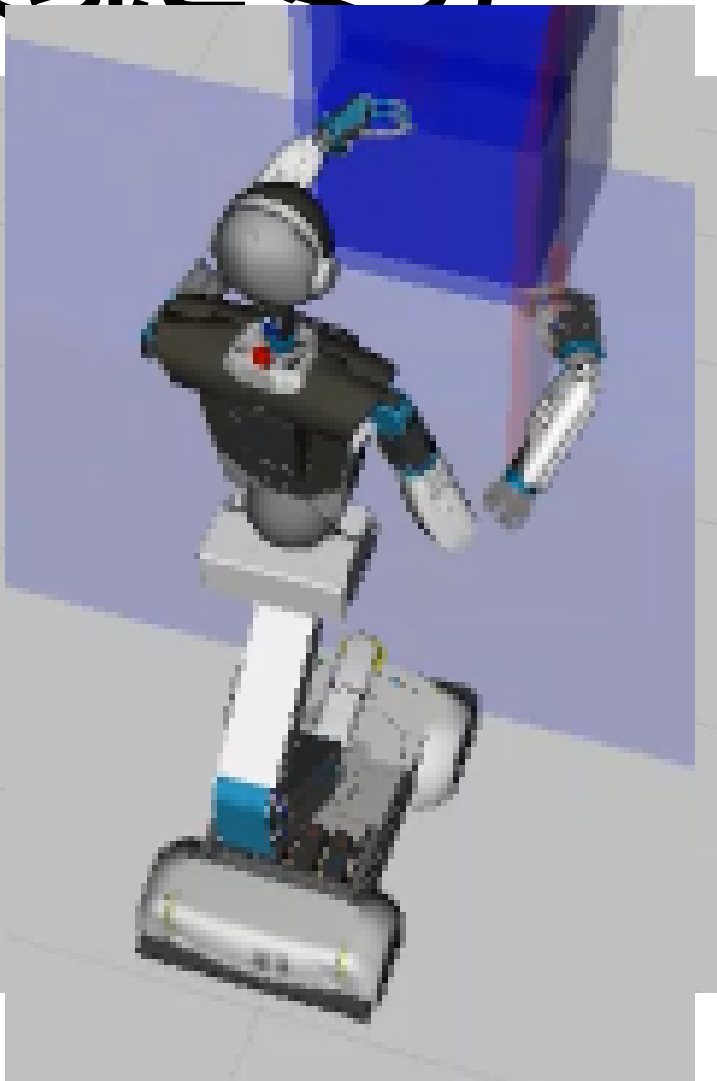




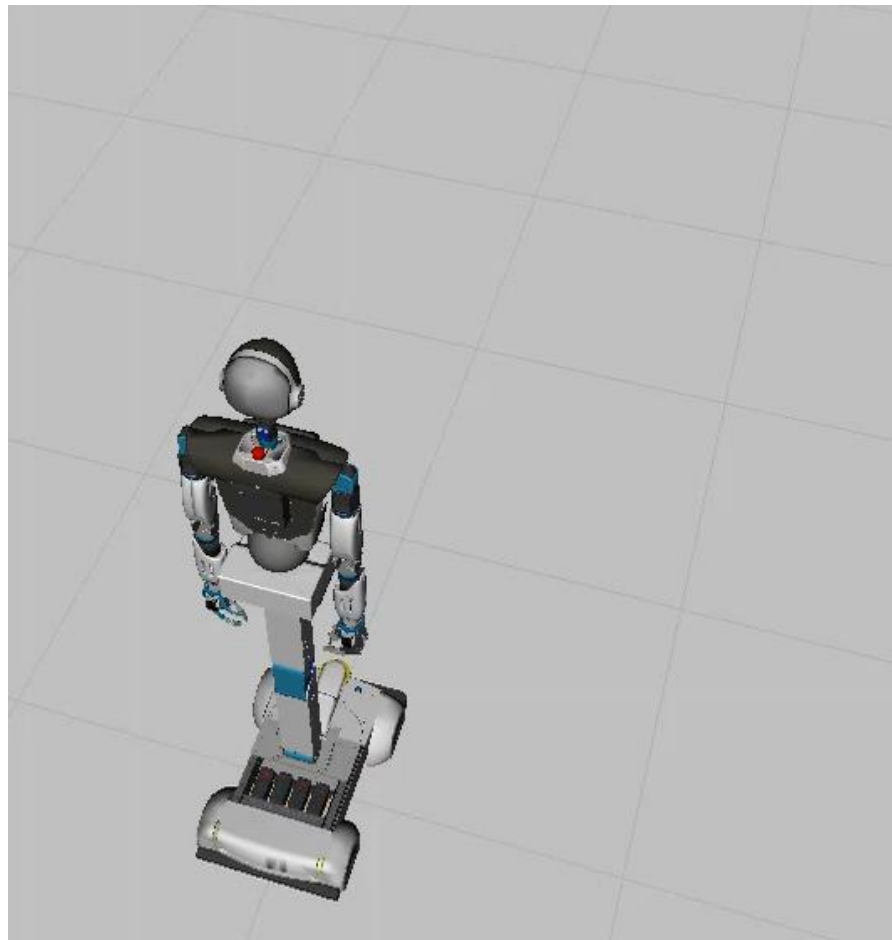
環境と姿勢の間の常識（目的因）

その目的のためこういった姿勢をとるのか

環境と姿勢



常識なし



常識あり

常識的な姿勢をとる

⇒ 予測ができ、
事故が少ない

ある姿勢は次の行動を
行う上で最適

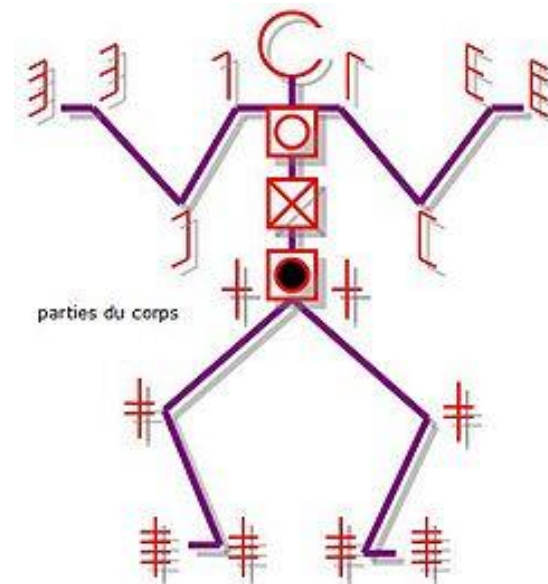
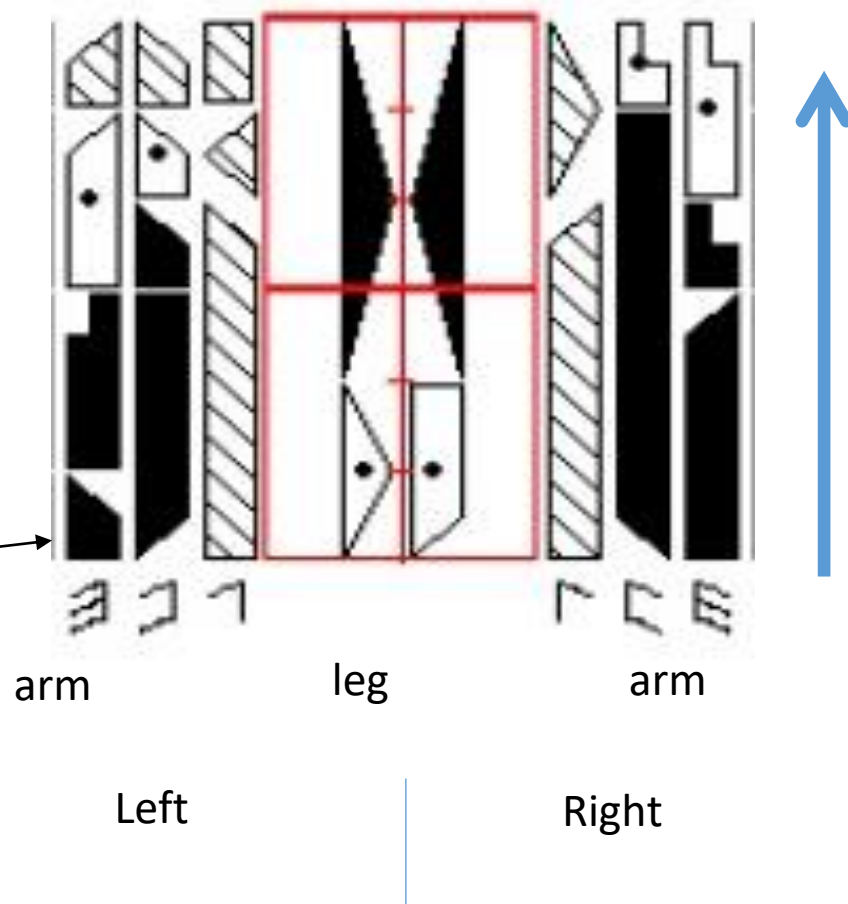
⇒ 隠れた計画

常識的姿勢を観察から

表現方法: ラバノーテーション

形: 方位角
色: 天頂角
長: 移動時間

Wrist pose
Azimuth: Left-forward
Zenith: Low
Length: Half beat



ラバノーテーション: 人間の感覚にそったデータ圧縮

時間離散化: 小休止

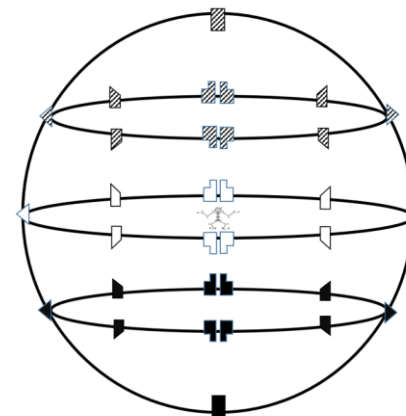
- 人間は重要なポーズで少し止まる



Key Frames (brief stop)

空間離散化: 8方位角、5天頂角

- 人間の離散化能力は 7 ± 2 /次元
(ミラーの理論)

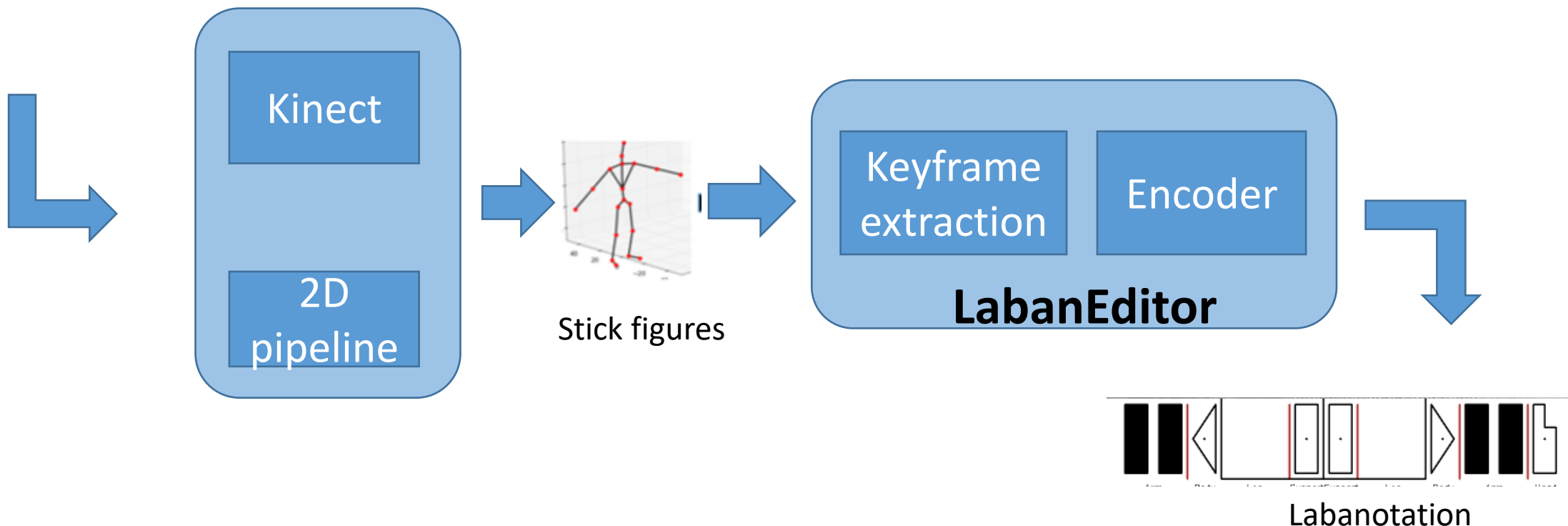


LabanSuite: ラバノーテーション抽出ソフト

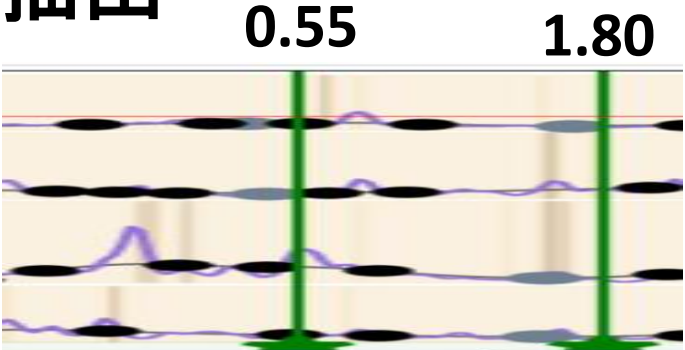
Applied Robotics Research Library



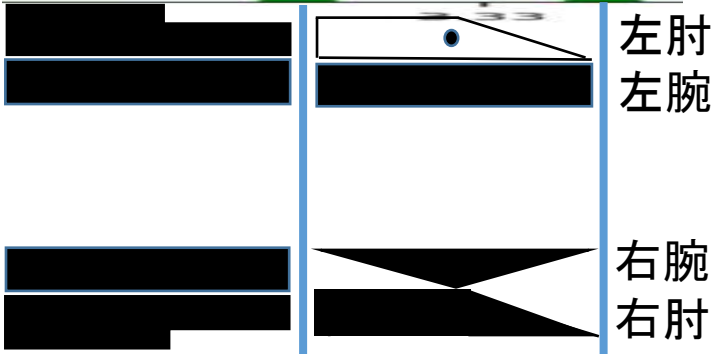
Motion sequence



LabanSuiteによるPick & Place動作抽出



LabanSuiteによる
キーフレーム
抽出



LabanSuiteによるラバノー
テーション

L:Idle
R:Pick

L:Idle
R:Place

Pick 動作

右腕: Place-low

右肘: Forward-low

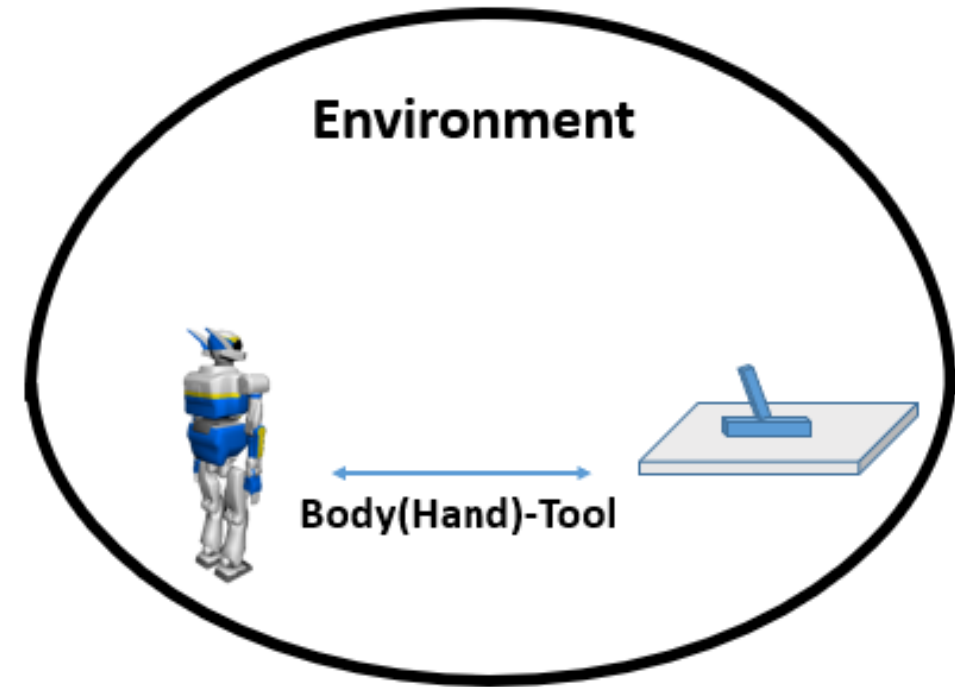


Place 動作

右腕: Right-low

右肘: Rightforward-low





体と道具の間の常識（目的因）

その目的のためにどういった把持をするか

手と道具の関係

- ・ 把持は動作の目的に応じて変化する

押すためには:



指すためには:





















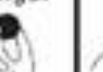









書くためには:

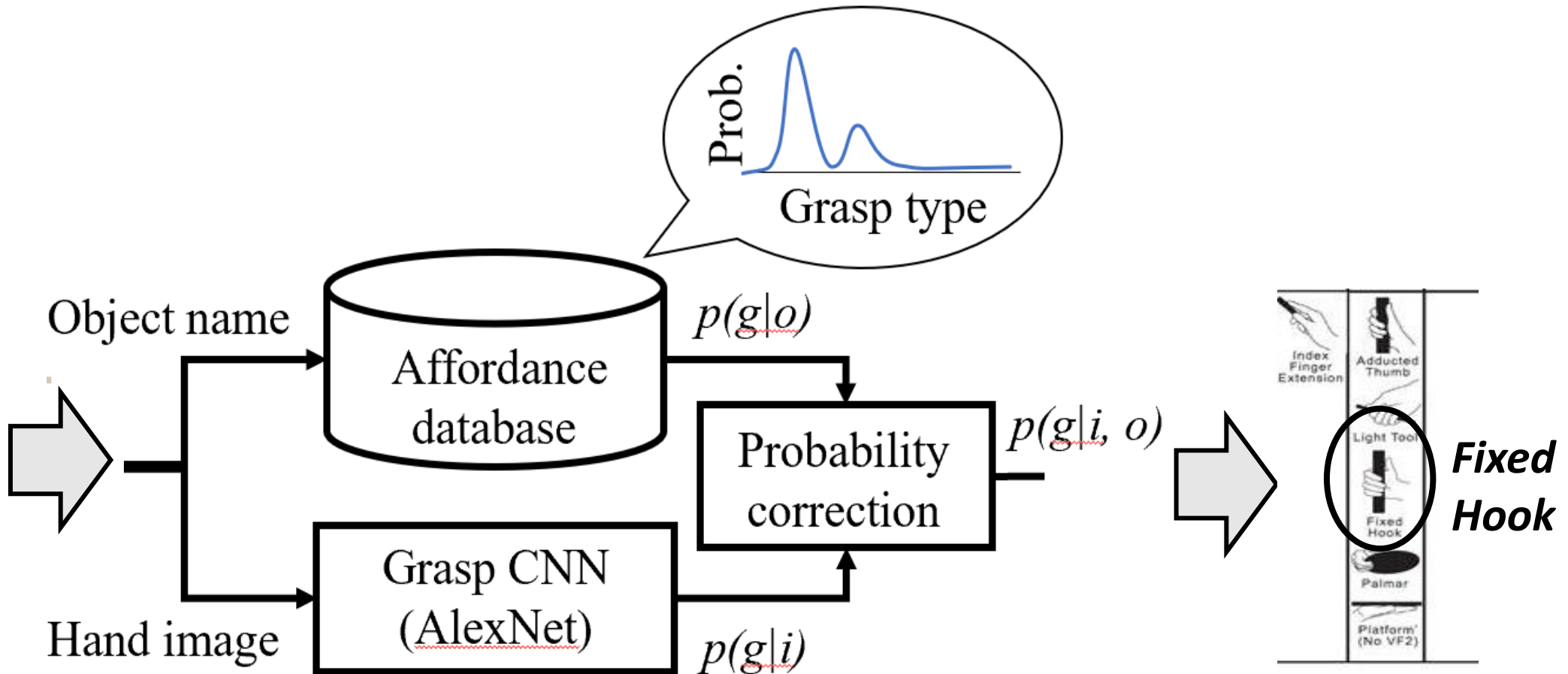


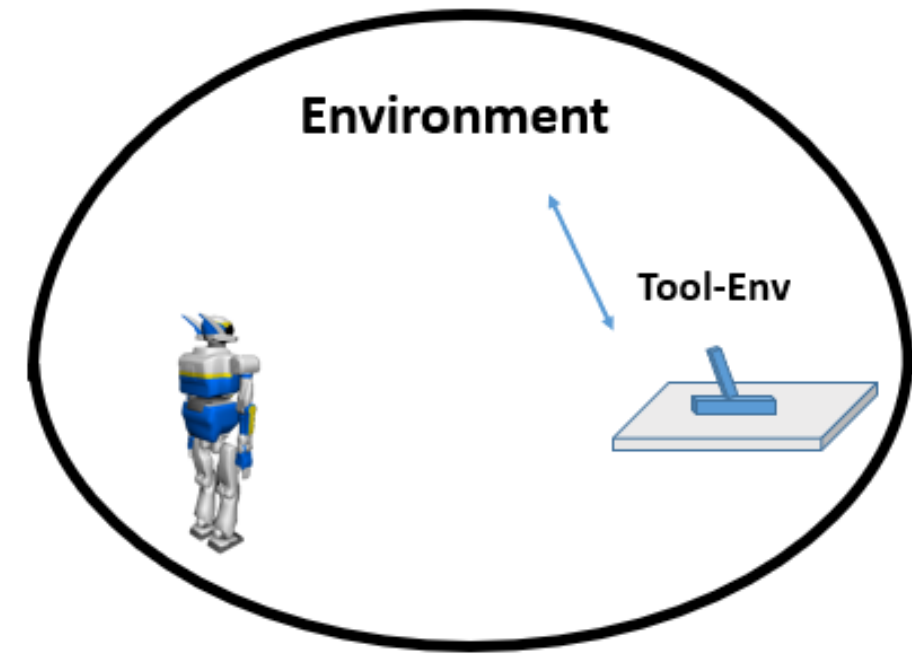
把持の分類

Felix 2016

Opp	Power						Intermediate			Precision				
	Palm		Pad				Side			Pad				Side
	3-5	2-5	2	2-3	2-4	2-5	2	3	3-4	2	2-3	2-4	2-5	3
VF2														
Thumb Abduction		 Large Diameter  Small Diameter  Medium Wrap  Power Disk  Power Sphere	 Ring	 Sphere-3 Finger	 Extension Type  Sphere-4 Finger	 Distal Type	 Adduction		 Tripod Variation	 Thumb-Index Finger  Tip Pinch  Inferior Pincer	 Thumb-2 Finger  Tripod	 Thumb-3 Finger  Quadpod	 Thumb-4 Finger  Precision Disk  Precision Sphere	 Writing Tripod
	 Index	 Adducted					 Lateral Pinch	 Lateral Tripod					 Precision Tripod	

CNN + affordanceによる把持認識





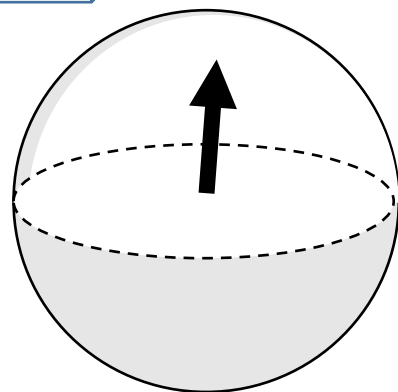
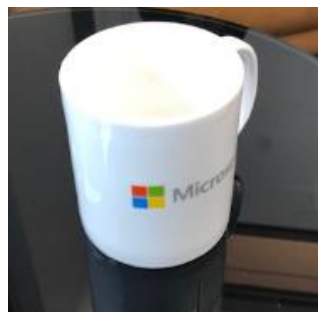
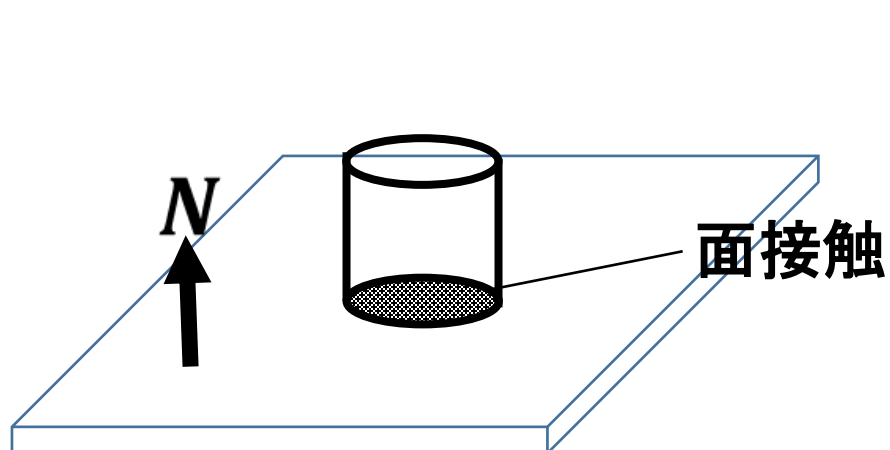
道具と環境の間の常識（目的因）

その目的のためは道具はどう動かされるべきか

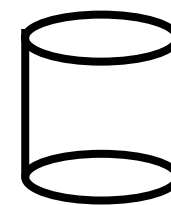
操作動作

操作とは面接触関係を変化させるもの

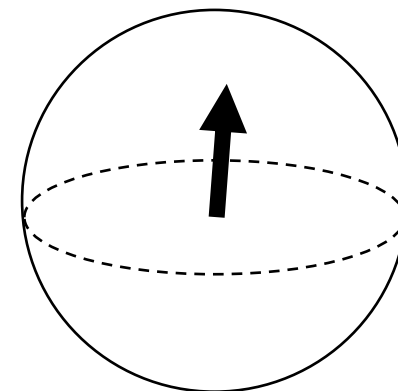
例: Pick-up 動作は、面接触をなくすもの



→
Pick-up 動作

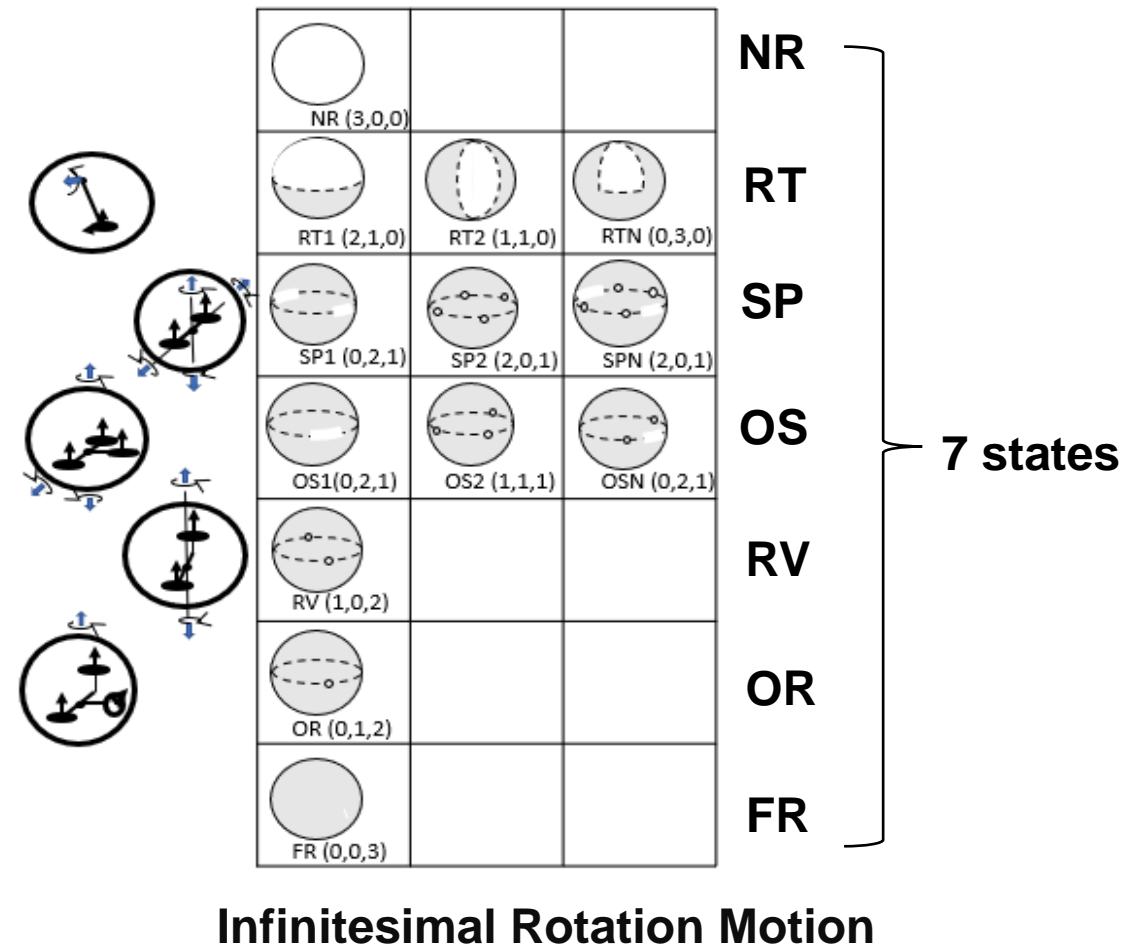
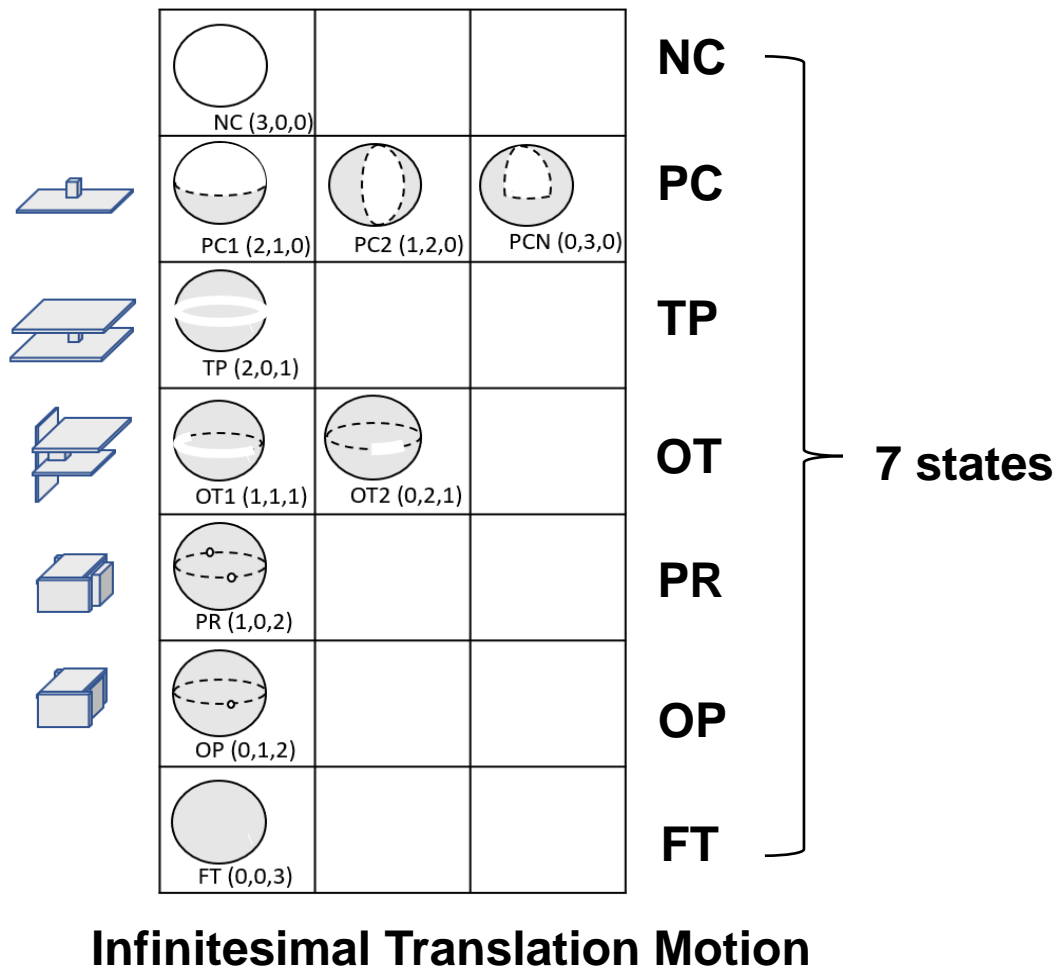


面接触なし



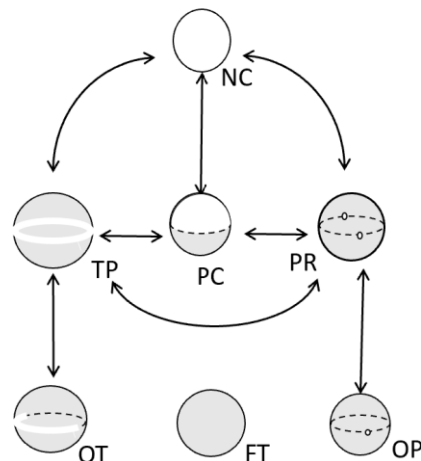
可能な状態

- Solution areas of simultaneous linear inequality equations
- Characterized using Kohn-Tucker theory and Screw theory



可能な遷移＝可能な操作

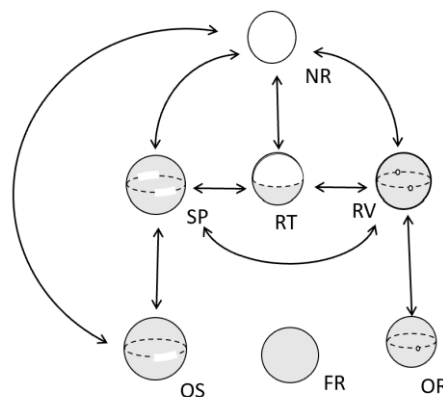
- Translation motion
7 states (start) x 7 states (end)
= 49 possible transitions



8 transitions

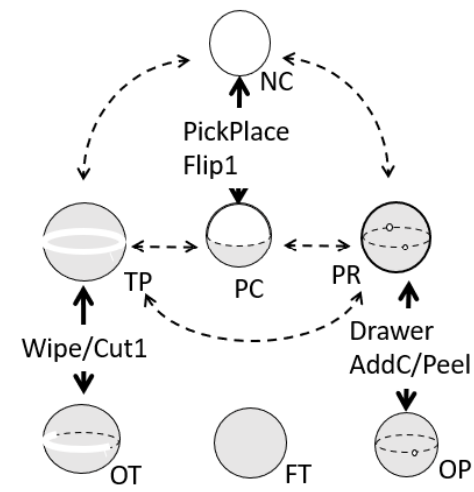
- Rotation motion
7 states (start) X 7 states (end)
= 49 possible transitions

Theoretical analysis

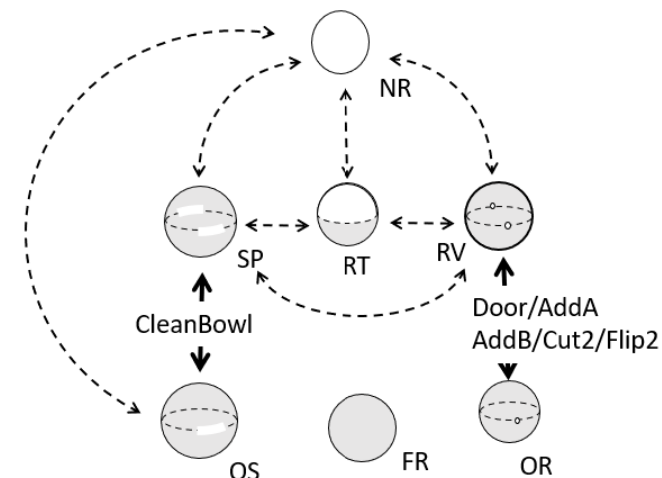


9 transitions

YouTube analysis



3 transitions
(2 physical & 3 semantic)



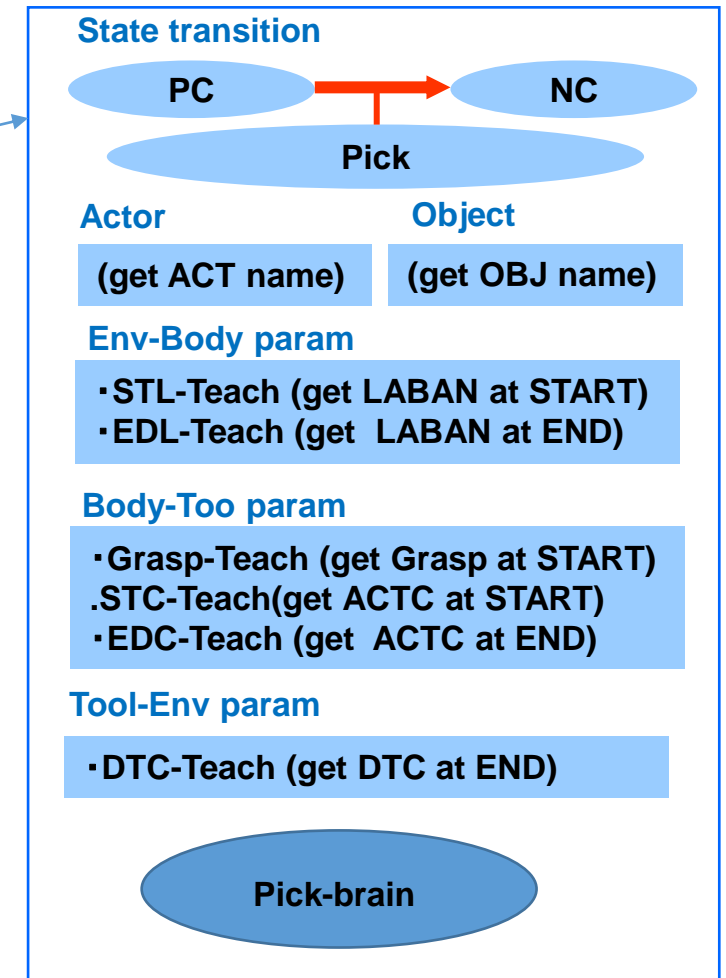
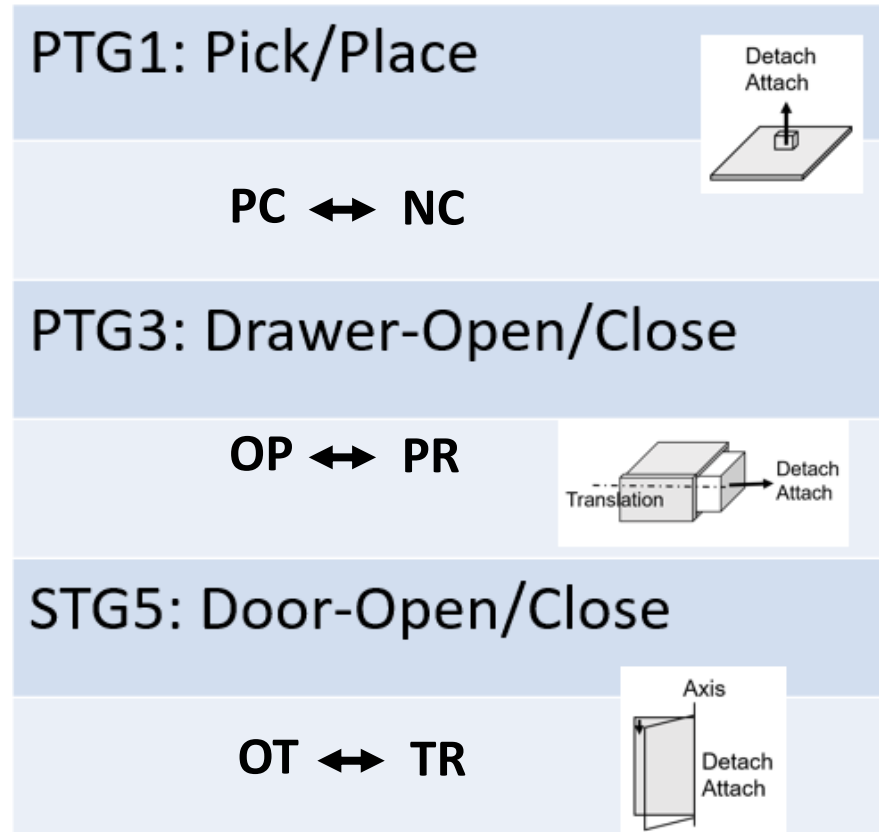
2 transitions
(1 physical & 2 semantic)

上限

物理的に可能な遷移

実際の発見された遷移

物理的接触による可能な操作（家庭内動作）



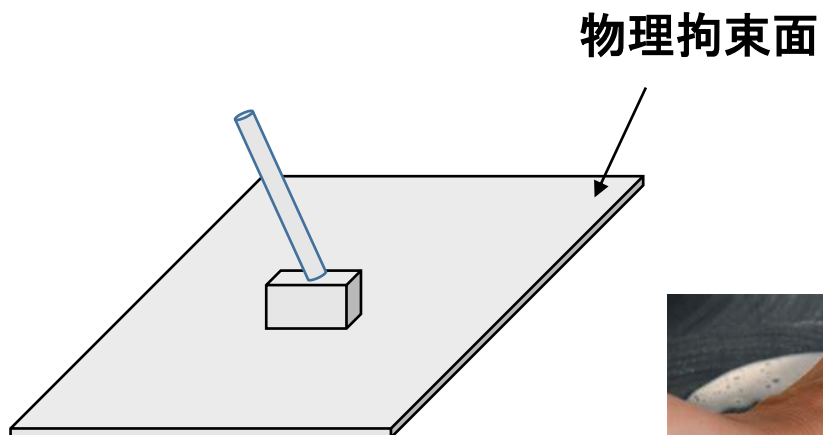
意味的操作

常識：面を拭くためには、接触を保持しなければならない

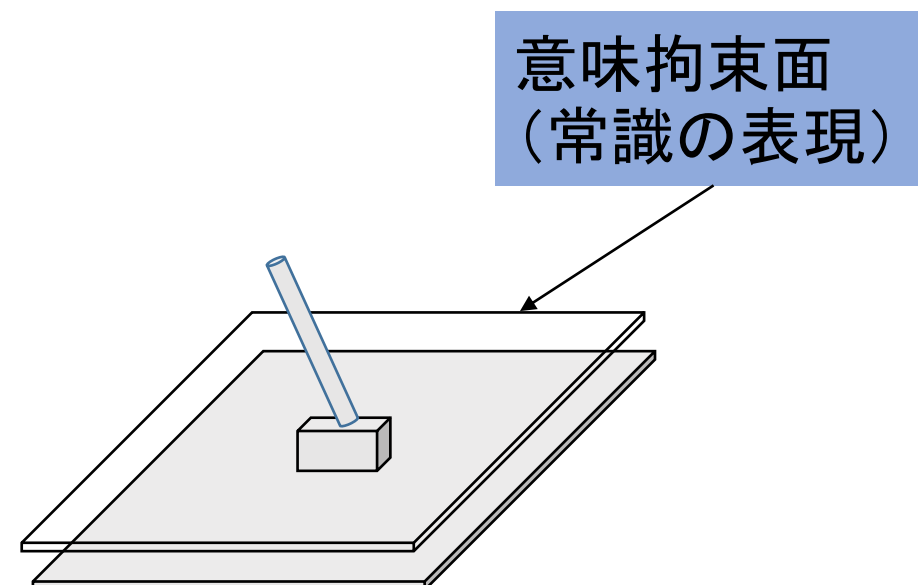


物理拘束と意味拘束

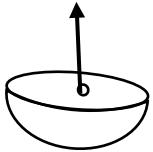

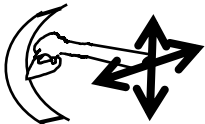

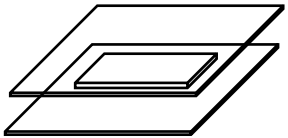

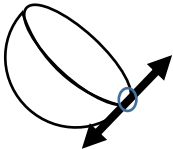
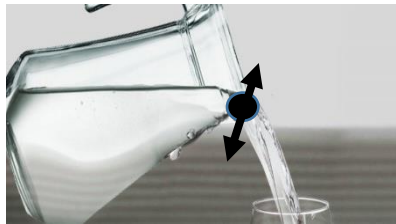
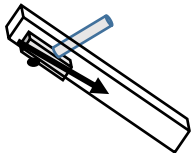

- 物理拘束: 1面接触一上方に動ける



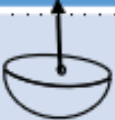
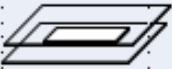



- 意味拘束: 拭くためには、面接触がなくなってしまうのではない



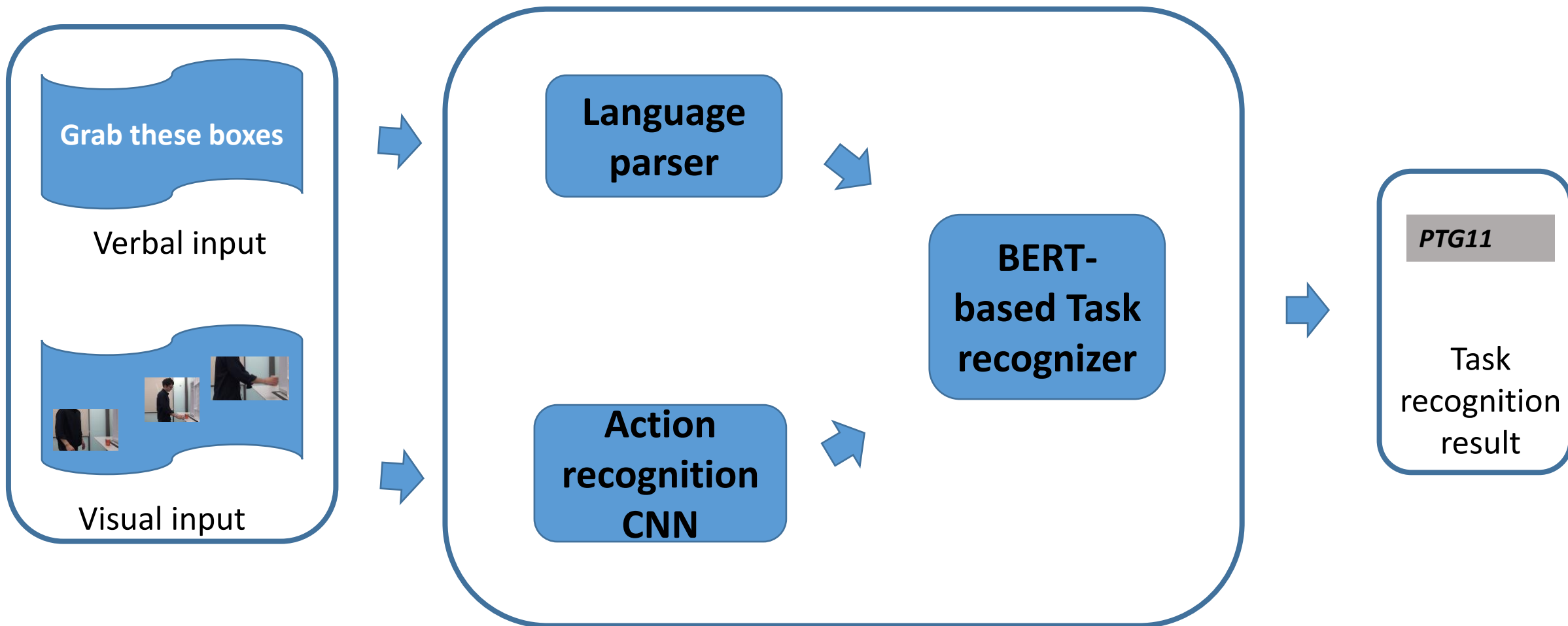
意味拘束面

Semantic Ping 		Semantic Sphere 	
Semantic Walls 		Semantic Hinge 	
Semantic Tube 			

意味的拘束からの操作動作

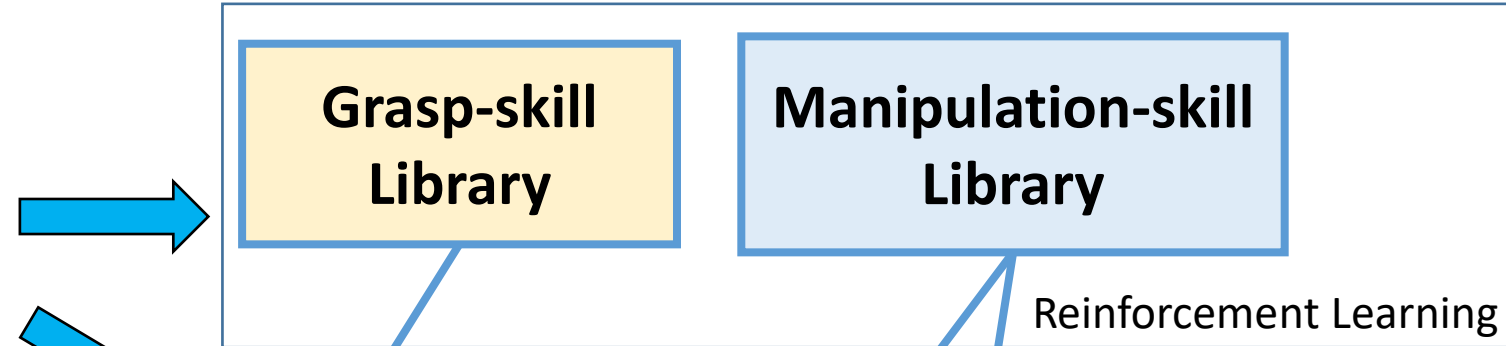
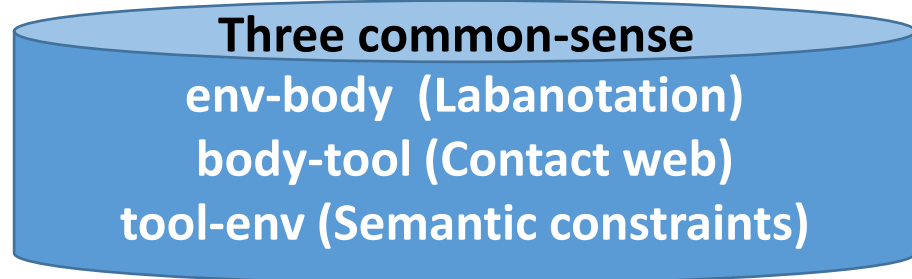
Constraints		Transitions	Task name & example
Semantic ping		PC ↔ NC	STG1: bring-carefully task
Semantic wall		OT ↔ TR	STG2: wiping task
Semantic tube		OP ↔ PR	STG3: peeling task
Semantic sphere		OS ↔ SP	STG4: bow-cleaning task
Semantic hinge		OR ↔ RV	STG5: pouring task

操作動作認識パイプライン

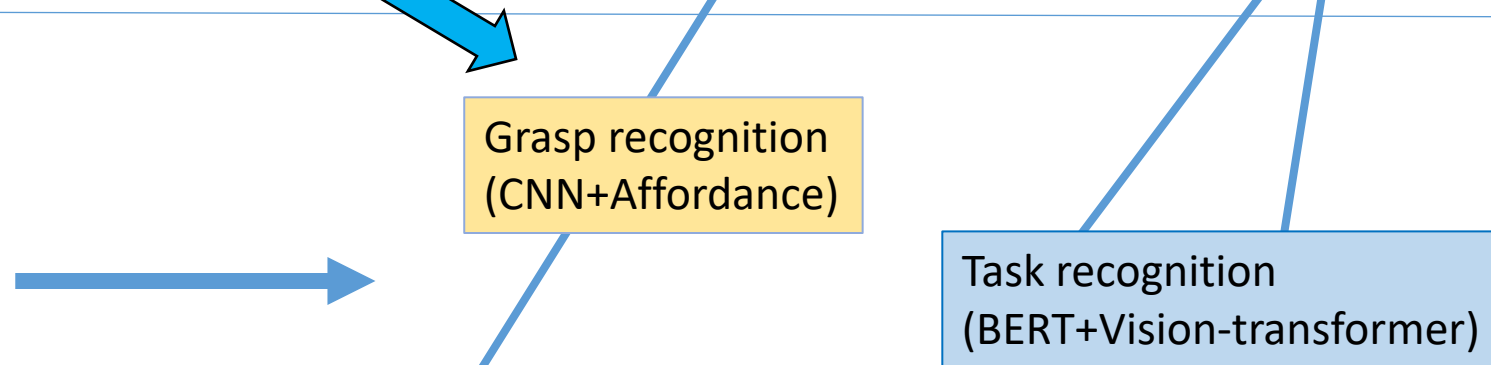


把持動作と操作動作ライブラリ

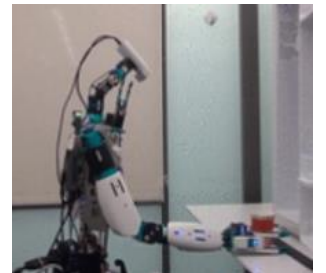
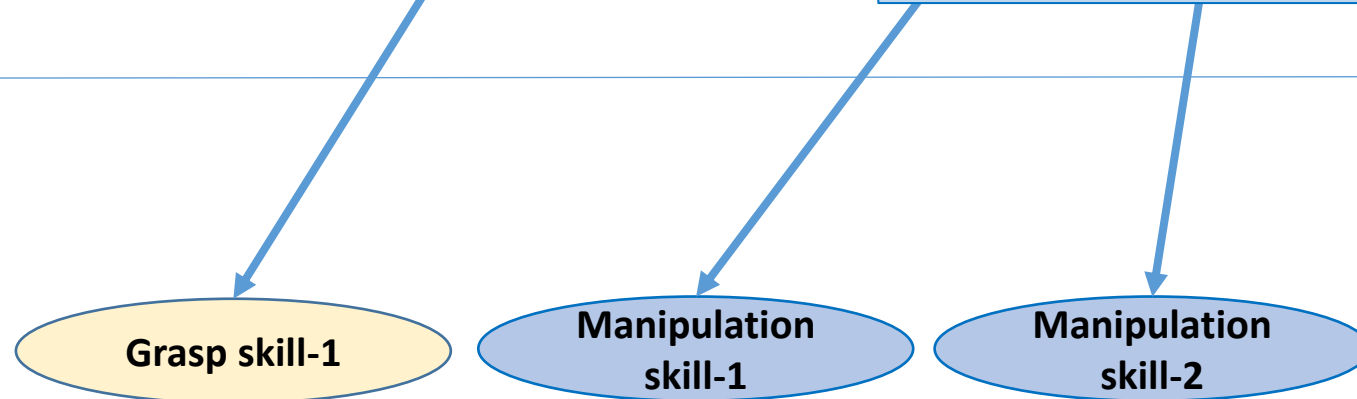
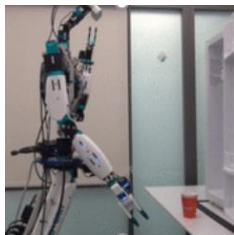
Design mode

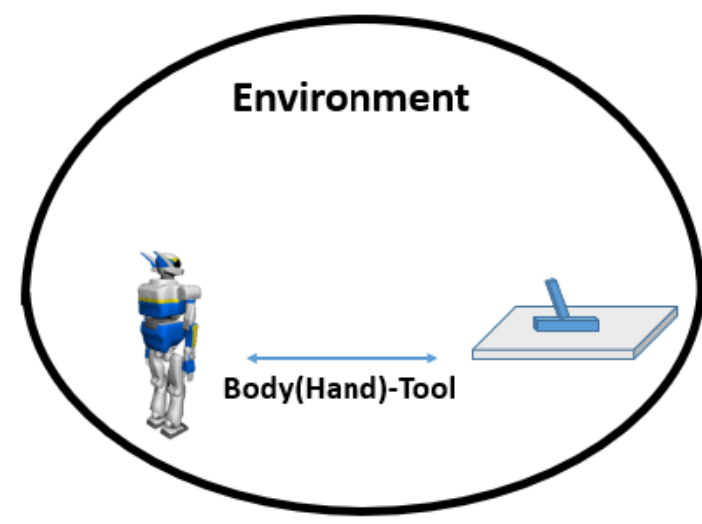


Encoding mode



Decoding mode

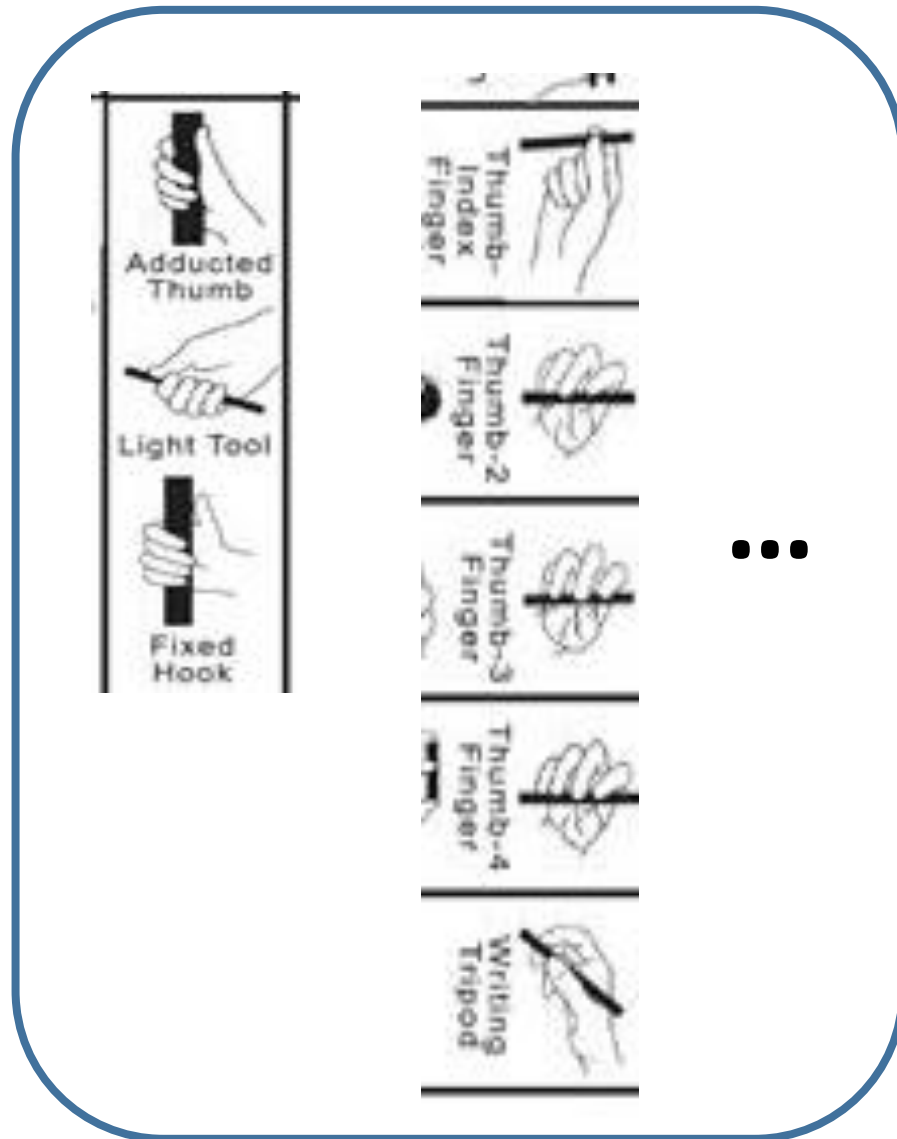




体と道具の間の常識にもとづく把持動作生成

再利用可能な把持動作ライブラリの生成

把持分類のクロージャ理論によるグループ化

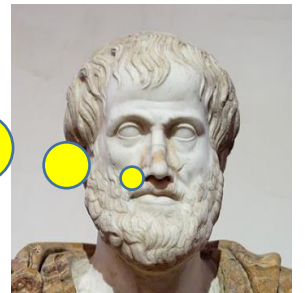


目的のためには冗長

Yoshikawa 2010

クロージャ理論

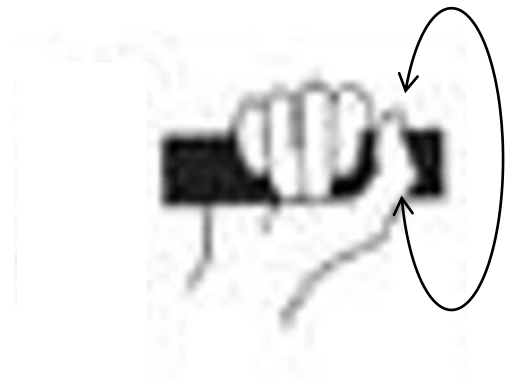
目的因：
なぜその把持を？



クロージャ理論(把持の目的)

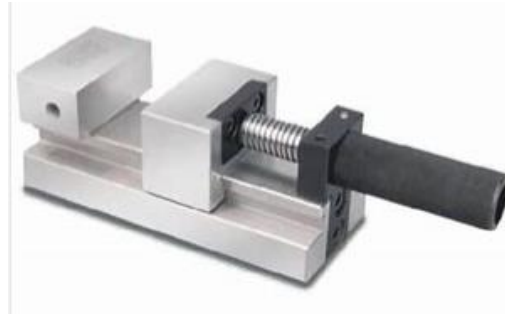
Passive form closure:

- 物体をその位置に保つ



Passive force closure:

- 物体に力をかけてどの方向にも動かなくする
















Active force closure:

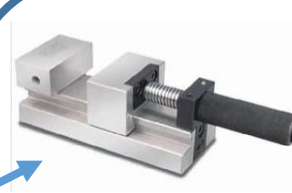
- 物体に力を掛けつつ動的に動かす



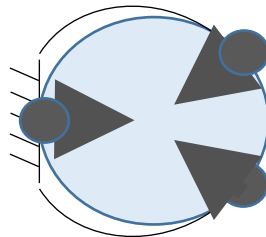
クロージャの実現としてのContact-webs

Large diameter		
Small diameter		
Medium wrap		
Power disk		
Power sphere		
Adducted thumb		
Medium wrap		
Extension type		
Fixed Hook		
Prismatic 2-finger		
Prismatic 3-finger		
Prismatic 4-finger		
Precision sphere		

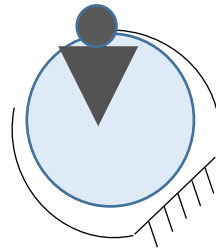
Grasp taxonomy



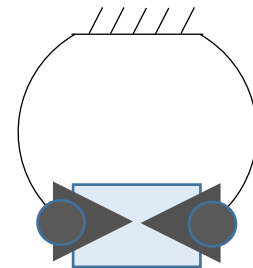
Passive force closure



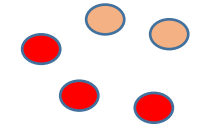
Passive form closure



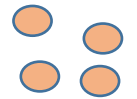
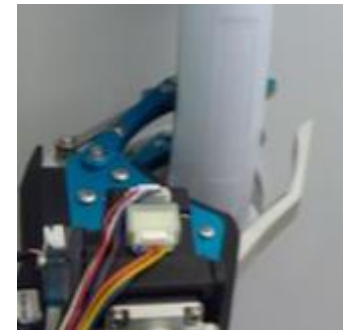
Active force Closure



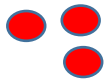
Closure theory



Passive force contact-web



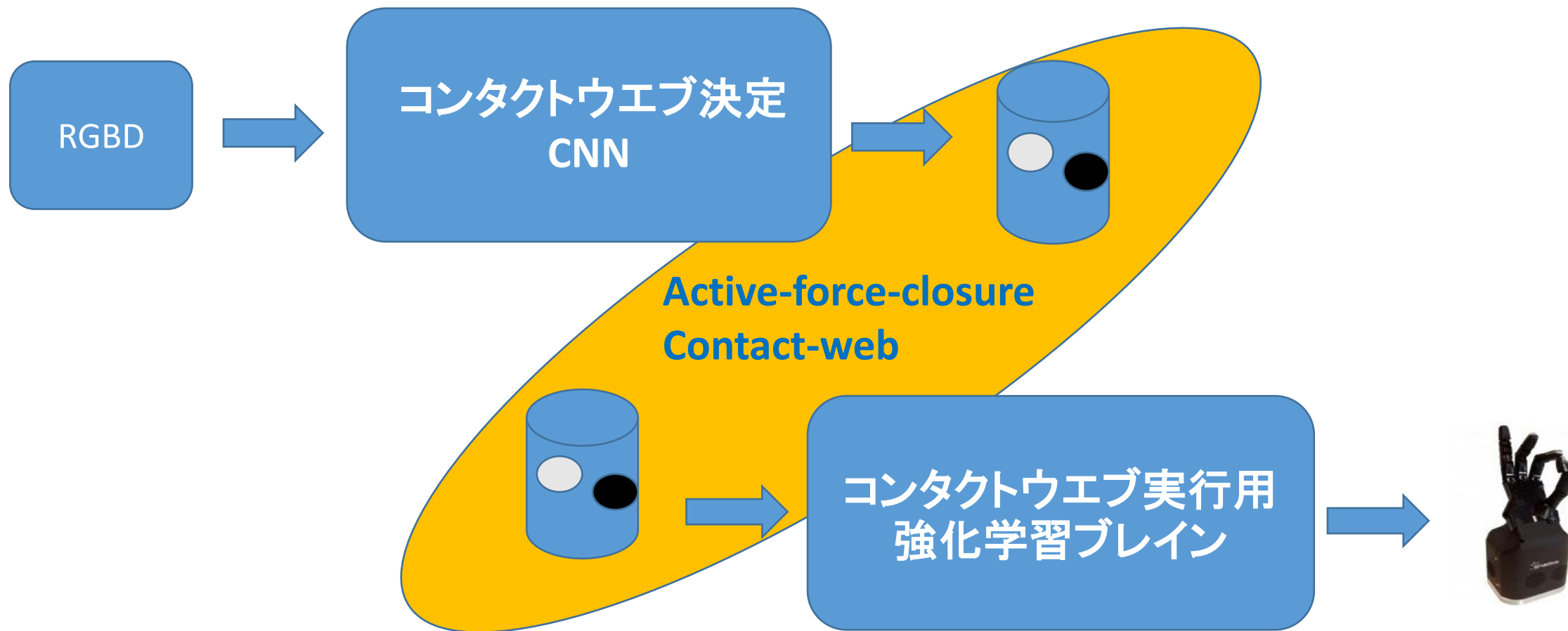
Passive form contact-web



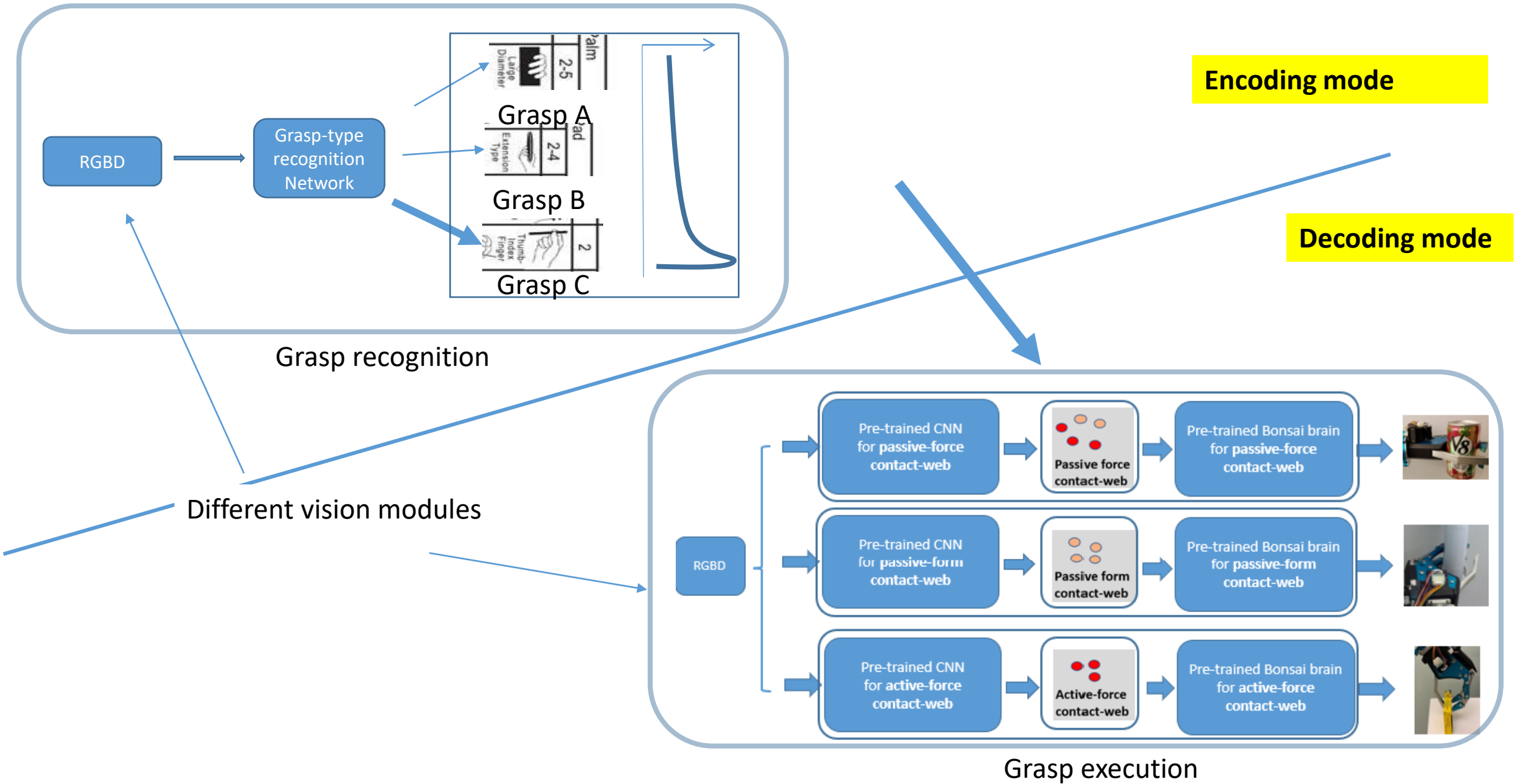
Active-force contact-web

Contact-web (SeedNoid)

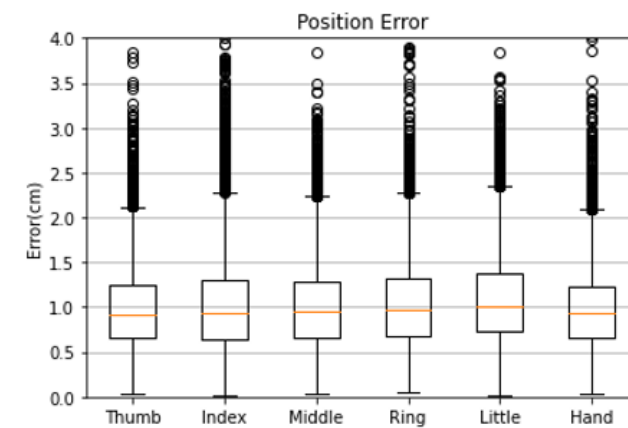
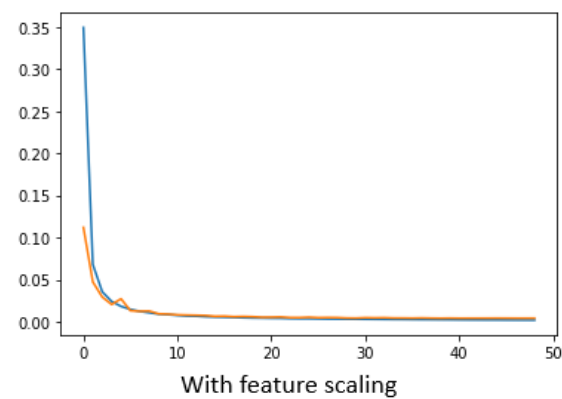
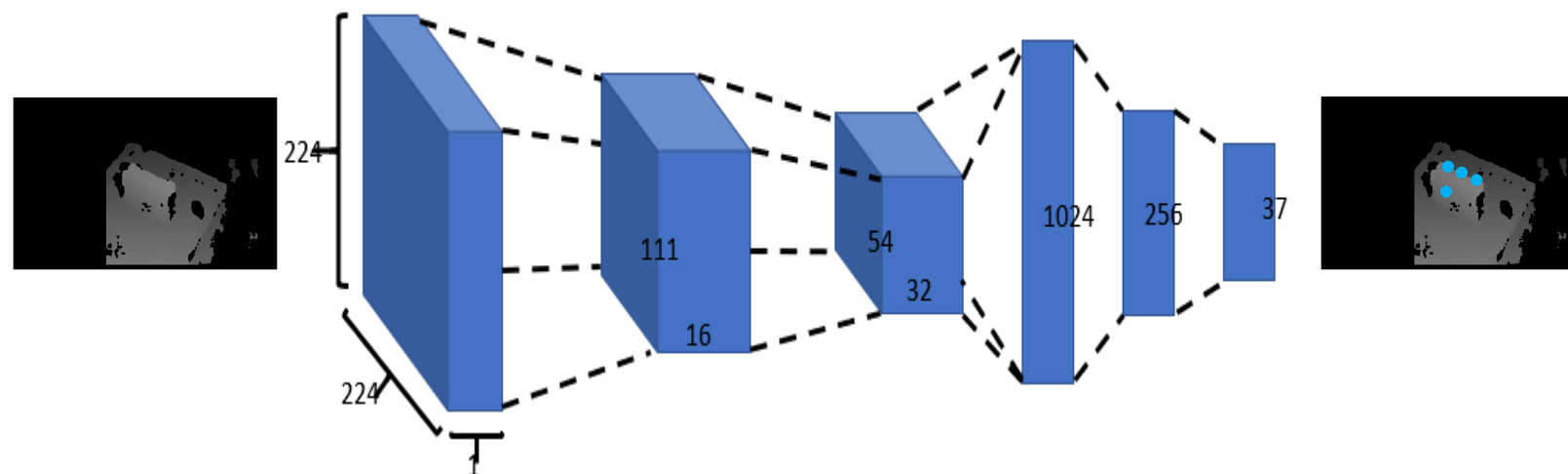
Contact-webベースの視覚動作パイプライン



ルーズコネクション



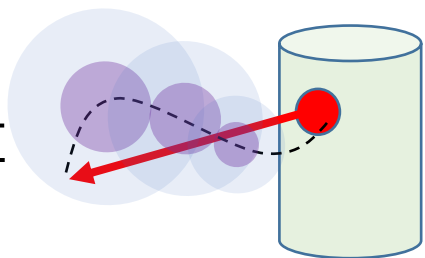
コンタクトWEB決定 CNN



Bonsaiによる強化学習

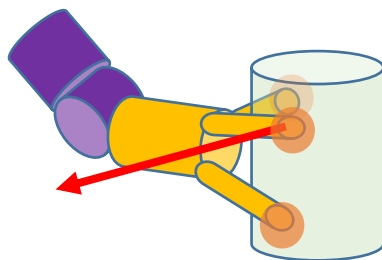
ヒント情報

- コンタクトウェブの位置



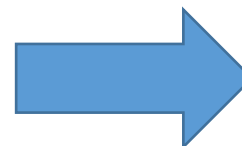
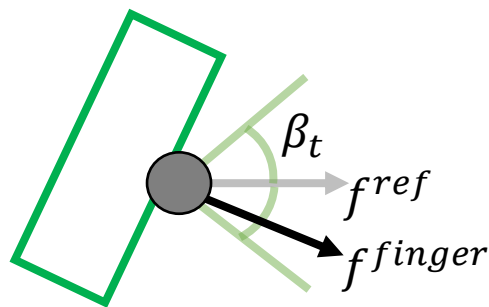
観測状態

- 指の位置
- カセンサの応答



報酬

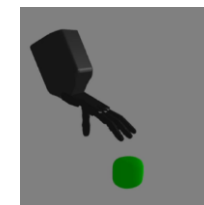
- 成功・失敗



強化学習システム

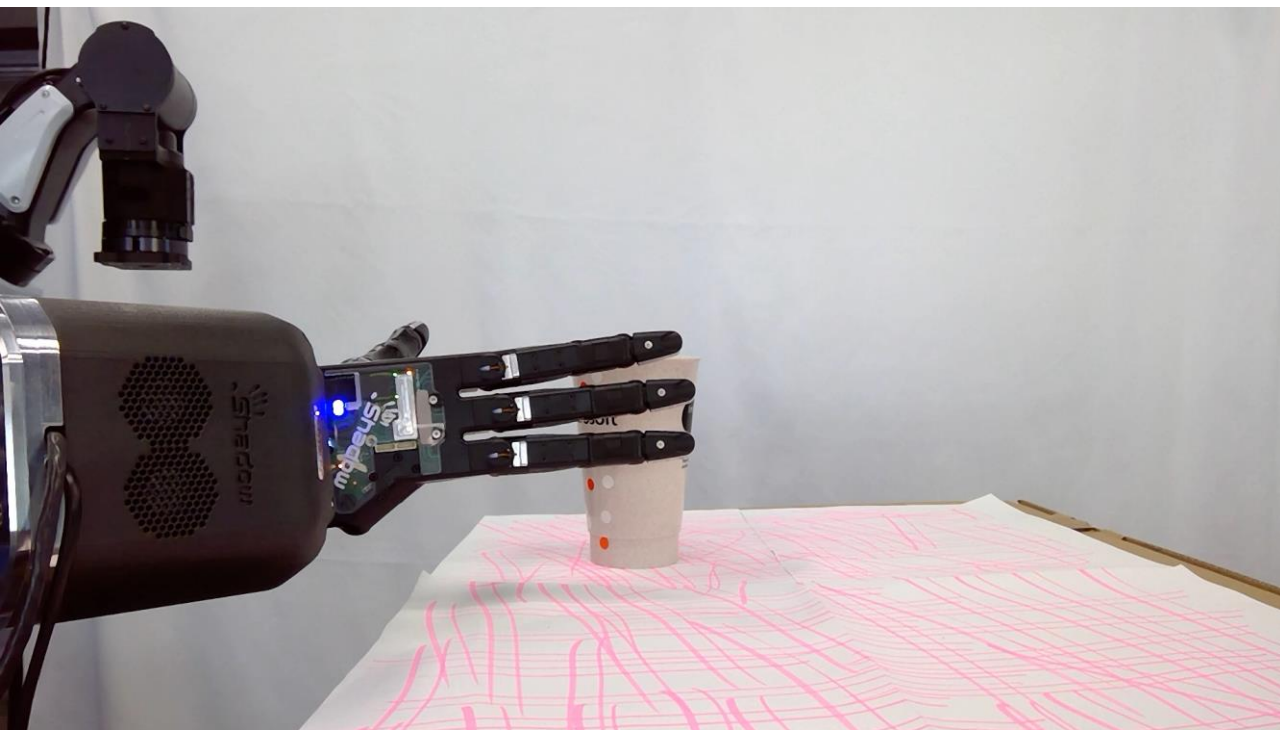


bonsai



Pybullet シミュレータ

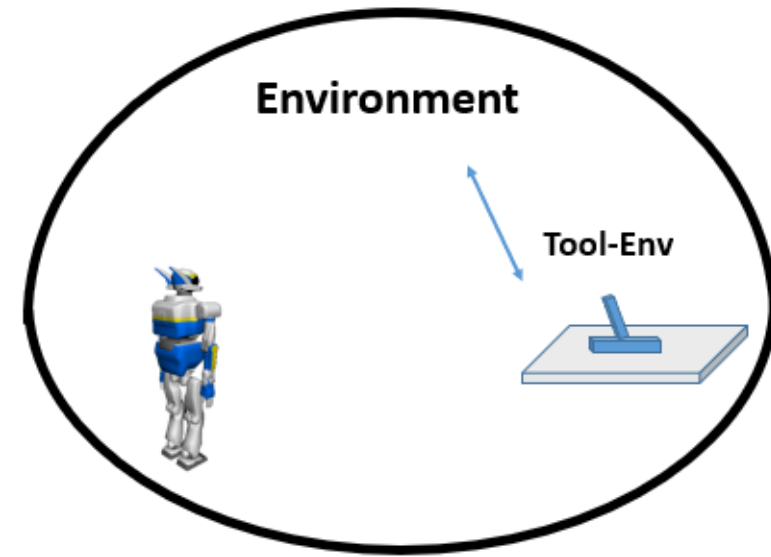
Passive force & Active force の実行



**Passive-force contact-web
(Shady hand)**



**Active-force contact-web
(Shady hand)**



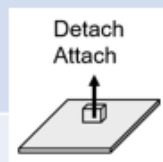
道具と環境の間の常識にもとづく操作動作生成

物理操作スキル

Physical Manipulations

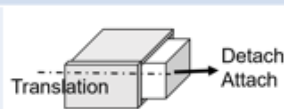
PTG1: Pick/Place

PC \leftrightarrow NC



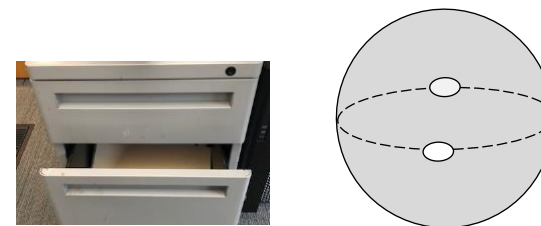
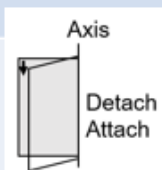
PTG3: Drawer-Open/Close

OP \leftrightarrow PR

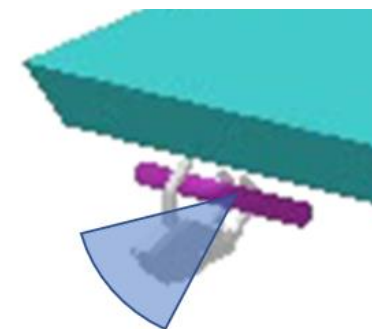


STG5: Door-Open/Close

OT \leftrightarrow TR



物理拘束



報酬関数

例: PTG5-引き出しスキル

ヒント情報

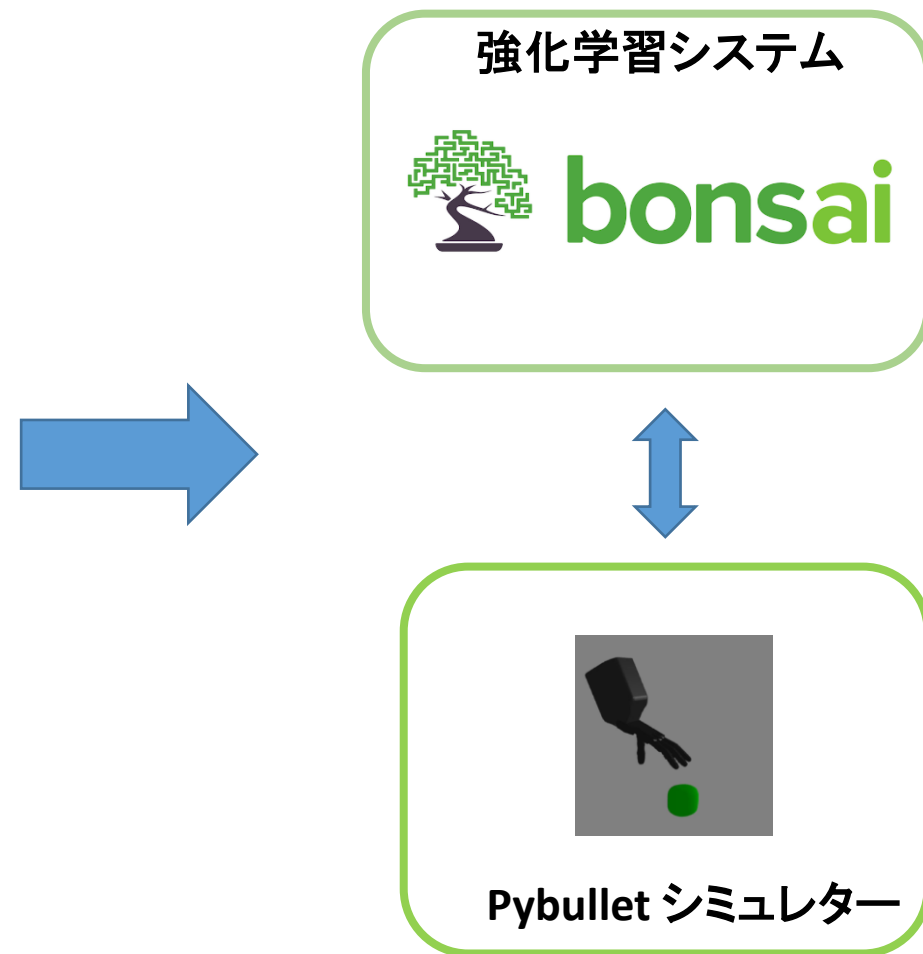
- 引き出す方向

観測情報

- 指の位置
- カセンサーの応答

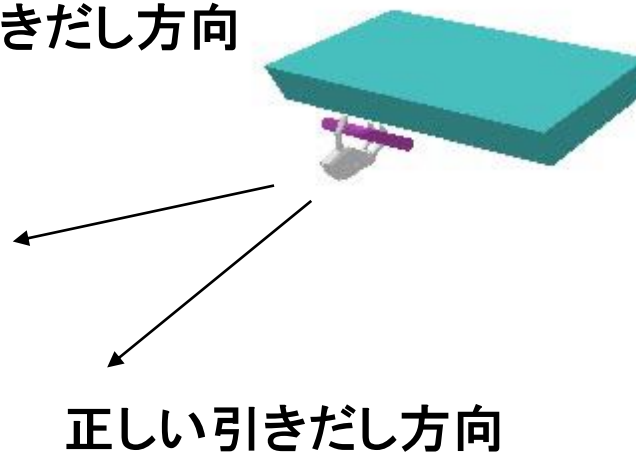
Reward

- 物理拘束からの力拘束
- 目的位置までの引き出し

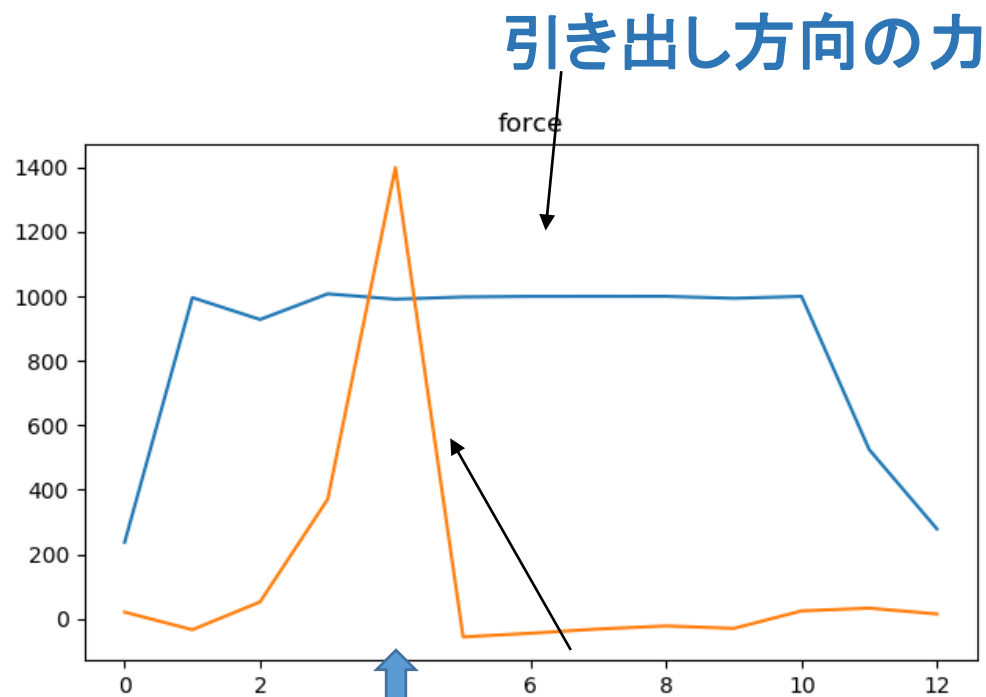


誤ったヒント情報下でのBonsai Brainの挙動

ヒント情報からの
誤った引きだし方向

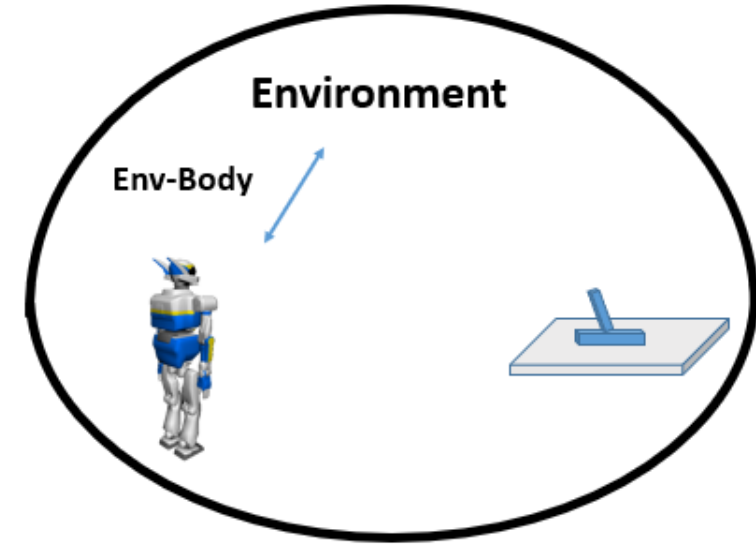


正しい引きだし方向



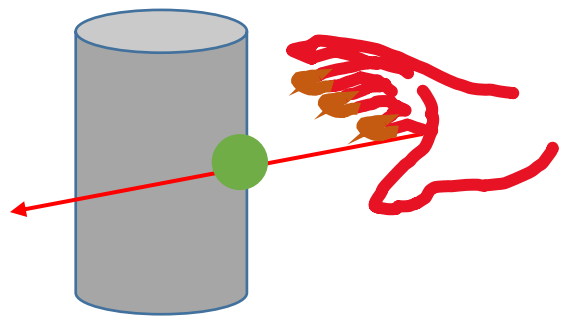
引き出し方向と垂直方向の力

引き出し方向の修正



環境と体の間の常識にもとづく動作生成

相対する2つのゴール



Bonsaiからの手の動き

+

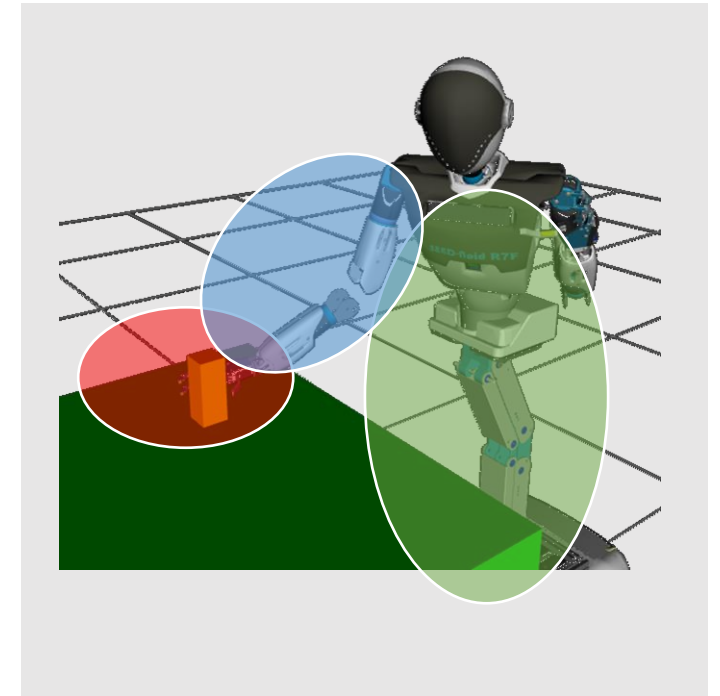


デモ (Labanotation) からの
腕と肘の姿勢

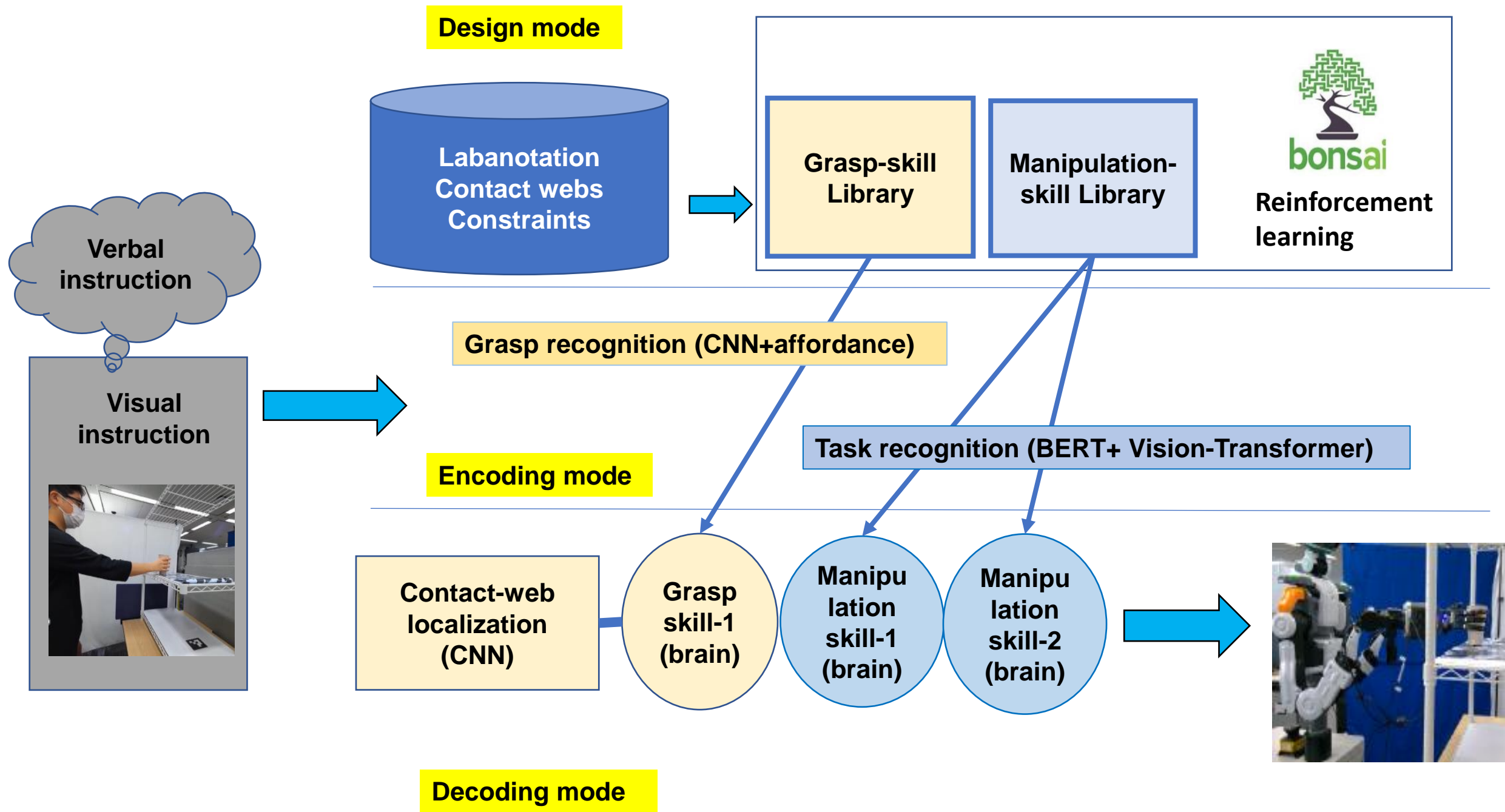
役割分担考慮アルゴリズム

The hand configuration is independent from upper/lower arm movements (Lacquaniti and Soechting 1982)

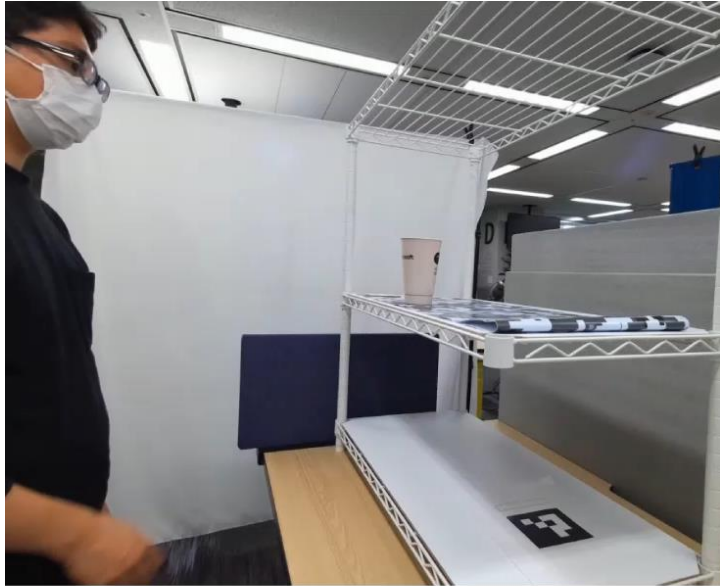
1. 手の動き: Bonsaiブレインより
2. 腕の動き: Labanotation & IK (demonstration parameter)
3. 体全体: 手と腕の調整



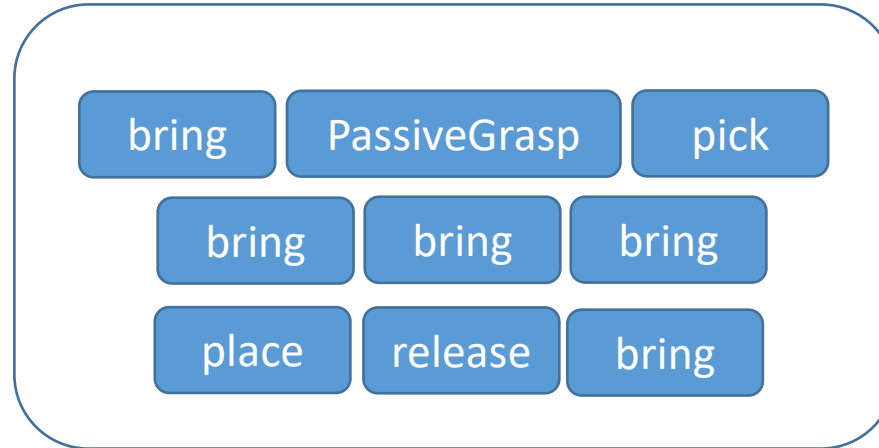
システム実行例



System in action (PassiveFG/Pick/Place)



**Encoding
task models**



**Task models
generated**



**Decoding
task models**

現在の仲間



Katsu
Ikeuchi



Jun
Takamatsu



Kazuhiro
Sasabuchi



Naoki
Wake



Atsushi
Kanehira



Daichi
Saito



Hiroshi
Teshima



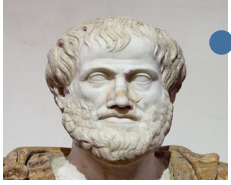
Machiko
Sato



Ryo
Hakoda

Redmond member

Shinagawa member



生命体は魂のために存在

- 現在のロボットは、みずから学びたいと思っているか
否： 思っていない。
単にプログラマーが学ばせているだけ

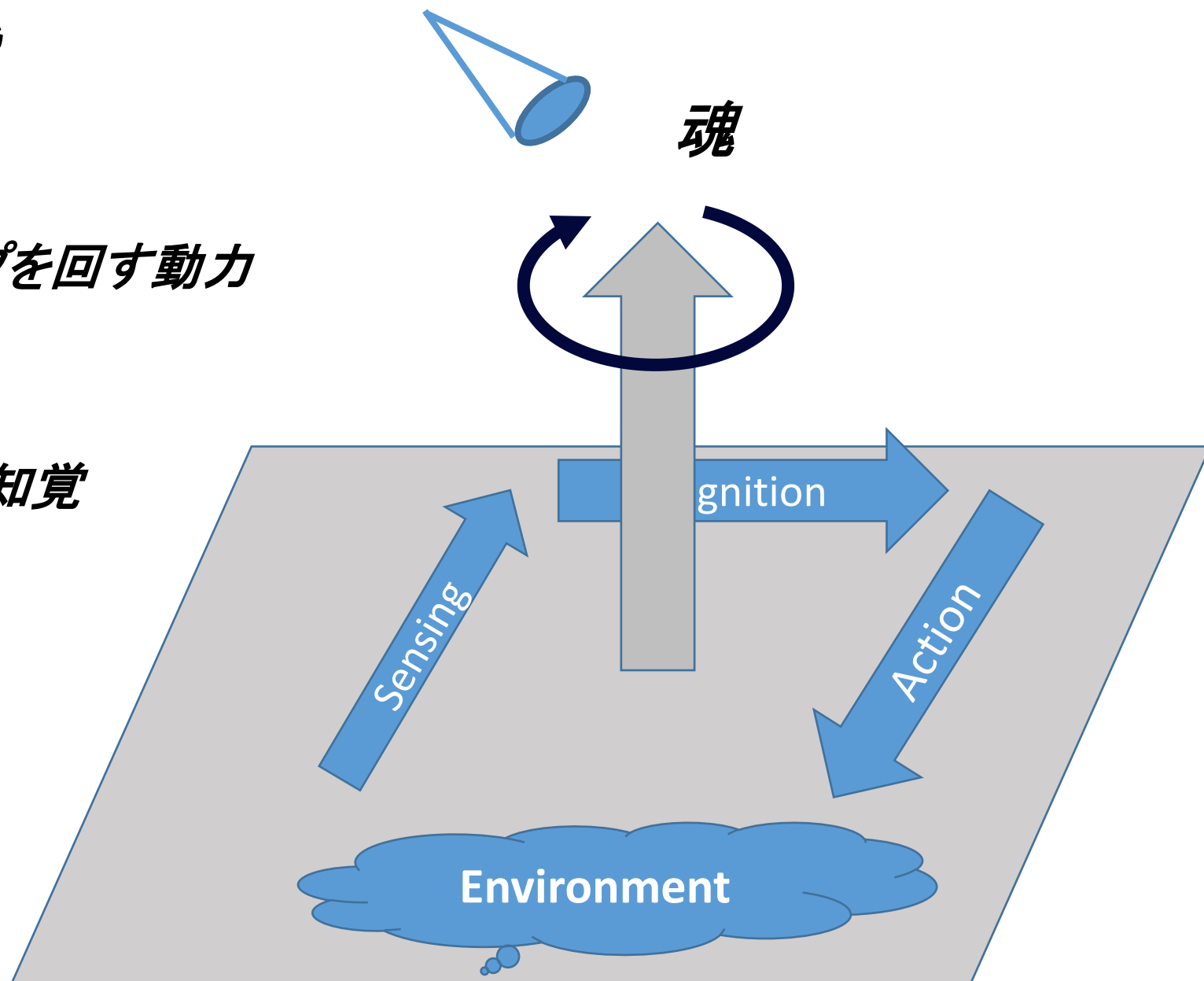
我おもうゆえに我あり



- 現在のロボットは、自我をもっているか
否： 自らを意識していない

疑似生命体の魂

- 生命体は魂のために存在:
魂はループ外にあり、ループを回す動力
- 我思うゆえに我あり:
魂は、回っていることを知る知覚



ロボット魂はサービスロボットに必要か？

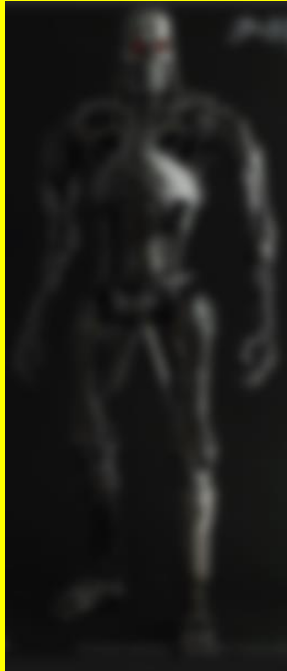
- ・作業仮説： ころのこもったパートナーのためには、“ロボット魂“が必要？

ロボットの究極的な目的が、疑似生命体をつくるなら、疑似魂をいれずには完成しない？

2つのタイプの疑似魂とアリストテレスの目的因

ターミネータ型目的因:

- 人と競争し
人を置き換える
- 完全自律自動
- Autonomous intelligence



ドラえもん型目的因:

- 人と協調し
人と共存する
- 物理的・知的に人間の活動を支える
- Augmented intelligence



Reddyとのディスカッション ー ドラえもん型ロボット

- ロボットは人間の友人 (ガーディアン) であるべき
- 人間が主、ロボットは従
- ロボット疑似生命体は、ロボット魂のために存在するが、
- このロボット魂は、*Autonomous Intelligence*でなく
*Augmented Intelligence*であるべき

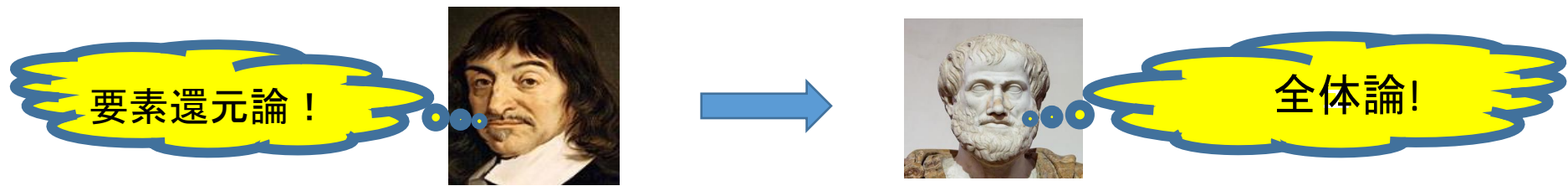
⇒ ドラえもん型魂



Reddy personal
communication 2017

第四楽章のまとめ

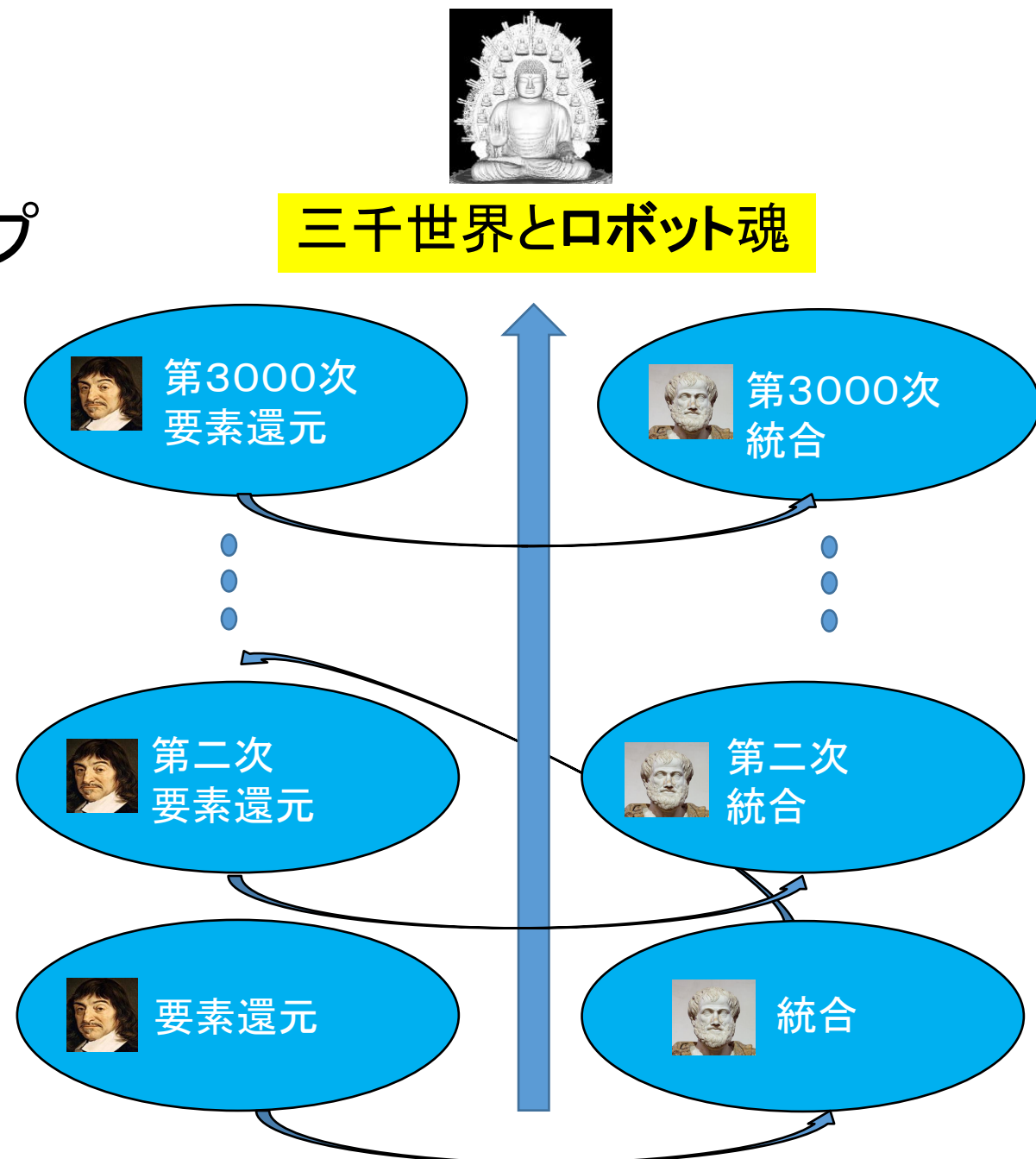
- パラダイムシフトの必要性
 - デカルト的要素還元の行き過ぎ：カンブリア爆発
 - アリストテレス的全体論による手法の統合



- 環境としての人間を考慮した目的因： 人間のためのAI
- 人間と競争するAI(Autonomous Intelligence)ではなく、
人間と協調共存するAI(Augmented Intelligence)へ

東洋的思考方

- 本来は、要素還元と統合のループ
正、反、合の繰り返し
- この三千世界のループから
 - ロボット魂の設計
 - 人間魂の理解



まとめ

- **第一楽章 (Andante) : 明るさ解析**
 - 環境に飛び込み、あせらず着実に、成果を出す
- **第二楽章 (Allegro) : e-Heritage**
 - 良き理解者を得て、速度を上げて、成果を出す
- **第三楽章 : (Vivace) 人まねロボット**
 - 良き仲間を得て、楽しく活発に、成果を出す
- **第四楽章 (Adagio) : パラダイムシフトとロボット魂**
 - 巨人の肩に立ち、緩やかに堂々と、成果を出す

この世で最も名誉ある職務を提供してくれる人よりも
私の自由な時間を妨げることなく好意をもって楽しませてくれる人の方を
つねにありがたいと思う

—— ルネ・デカルト 方法序説

THANKS