

招待講演：Web 上における大規模データの 効率的構築および有効利用に関する試み

川 崎 洋^{†1} 河 合 由起子^{†2} 古 川 亮^{†3}
小野 晋太郎^{†4} 池 内 克 史^{†4}

近年、Web 上に蓄積されるデータ量が加速度的に増加・大規模化する中、それらを効率良く構造化し、利活用するためのサービスが求められている。例えば、Google Street View は、世界規模で主要道路における全方位画像を数メートル間隔で蓄えており、ユーザは地図上で見たい地点をクリックするだけでその画像にアクセスすることができる。我々は、このようなサービスが成功する条件として、大規模データが効率よく構築できること、および、構築した大規模データを効果的に利用できることが必須と考え研究を進めてきた。ところが、実世界地図データや 3 次元データ等の大量データでは、両方の条件を満たすようなサービスがなく、普及しているとはいえない。本講演では、このような視点に基づき、Web の大量データを用いたサービスが成功する条件について考察すると同時に、我々が上記データを対象として、これまで取り組んできた研究事例について紹介する。

Invited Talk: Efficient construction and utilization of Massive data on the Web

HIROSHI KAWASAKI,^{†1} YUKIKO KAWAI,^{†2} RYO FURUKAWA,^{†3}
SHINTARO ONO^{†4} and KATSUSHI IKEUCHI^{†4}

Recently, efficient construction and utilization of massive data on the Web is strongly required. The authors have been researched on that area more than decade especially with geographic data, car-mounted video data and 3D data of real objects. In this talk, several solutions with actual applications and future direction of the projects are shown.

^{†1} 鹿児島大学

Kagoshima University

^{†2} 京都産業大学

Kyoto Sangyo University

^{†3} 広島市立大学

Hiroshima City University

^{†4} 東京大学

Institute of Industrial Research, University of Tokyo

1. はじめに

我々の研究グループでは、Web 上の大量データには、以下のような大きな 2 つの課題があると考え、これまでその解決に取り組んできた。

A 効率的な構築方法^{*1}（大量データの集め方および構造化方法など）

B 利用方法（構造化データの効果的利用が可能なインターフェースなど）

上記課題について、これまで数多くの解決手法が提案されてきた。例えば、構築方法に関しては、これまではテキストベースだったものが、画像となり、今ではビデオのような巨大なデータも、効率よく集められ構造化されるようになってきている。また、利用方法に関しては、単純な検索サービスに始まり、最近では Google Street View のように、Web 上で地球の裏側をクリック一つで自由に歩き回れるようにまでなっている。これら 2 つの課題は、あるサービス (B) 実現のために大量データが必要となり、新しい構築手法 (A) が提案されたり、逆に既存の大量のデータ (A) に目をつけて、新たなサービス (B) が提案されるなど、両者が密接に結びつけられて研究されてきた。例えば、これまで Web 上で大量データを用いて成功したサービスとしては、Google Street View、画像検索、Photo-Synth といったものがあげられる。それらを A（構築法）と B（利用法）に分類したものを表 1 に示す。このように Web 上に大量データをうまく集める方法と、それを上手に利用する手法の 2 つが、高いレベルでバランス良く機能していることが、大量データを利用したサービス成功の秘訣といえる。

以上の視点から、我々は、「実世界地図データ」と「3D データ」という本質的に大規模である 2 つを対象に、上記の課題を解消し、誰もが使いやすい Web サービスの実現を目指して幾つかのチャレンジをしてきた。以下では、それぞれのテーマについて、具体的な事例および、それらを通して明らかとなった課題、今後の展望などについて紹介する。

表 1 過去の大量データを用いたサービス例

「サービス名」	「集め方・構築方法」	「利用方法・解析方法」
Google Street View	全方位車載カメラ＋電子地図	GPS, GIS
画像検索	Web 上の画像など	特徴抽出＋PageRank＋検索
Photo-Synth	ユーザーアップロード	Bundle adjustment＋クリックナビゲーション

^{*1} Web にあるデータ量のみで考えると、既に大量と言えるが、構造化されていない状況においては利用方法が限定されるため、ここではデータの構造化を含めた構築を指している

2. Web 利用のための実世界地図データの効率的構築

実世界のデータを構造化して地理データに関連づけ、Web 上などで利用するという考え方は、GIS や ITS などに関連した実用性の高いサービスに展開できるとして注目を集めてきた。ここでは都市シーンを仮想空間上に構築する試みに関して我々が取り組んできた事例を紹介する。

表 2 地理データを用いた Web システムの提案手法

「サービス名」	「集め方・構築方法」	「利用方法・解析方法」
仮想化都市構築	車載センサー+電子地図	時空間解析, DP +バーチャルリアリティ
Video-Synth	全方位車載カメラ映像+ Web 上のビデオ	SFM, ビデオの類似性+ビデオ合成

2.1 車載センサによる 3 次元都市構築

2 次元・3 次元の都市シーンを仮想空間上に構築する試みはこれまでも何度も行われてきたが、Google Street View に代表される成功例の一方、失敗した試みも少なくない。分かりやすい理由としては、見た目の現実感の低さや未成熟なインタフェースもさることながら、より本質的には、仮想空間と現実空間がリンクしていないことではないだろうか。

我々は、仮想空間を表すデータとしてインターネットの普及期から Web 上で扱われてきた電子地図などの地物データを、現実空間を表すデータとして車載カメラやレンジセンサのスキャンデータを考え、特徴に着目して両者を適切に組み合わせることを試みた。例えば、1) では、電子地図の建物モデルに対し、計測車両で得たレンジデータから建物のエッジを抽出して両者を DP により照合させることで、地面からの視点に基づいた詳細な表面形状を付加させるとともに、建物の高さを適切に設定した。副次的な効果として照合の結果から建物群に対する車両の位置履歴が得られることも特徴である。2) では、車載全方位カメラの映像から建物列を認識し、同様に照合することで建物の面状にテクスチャを付加している。建物列を安定して認識するため、画像中で空部分との境界を表すエッジに着目した時空間画像である時系列高さ画像を考案し、これを照合の特徴量として効果的に活用している。図 1 では、正しく境界が検知され、建物ポリゴンにテクスチャがきれいに付与されていることが確認できる。

また、現実空間のスキャンデータを構造化するためには、特定のデータ群をレイヤーとして分離できるようにする必要がある。4) では、シーンが壁面などの平面群の集合で近似的に構成されるとの仮定に基づいて、全方位映像の時空間解析によりデプスの異なる部分を分離・補完している。これにより、例えば街路において、建物壁の手前に存在する歩行者や

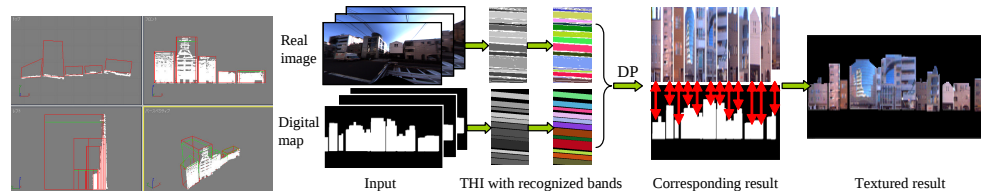


図 1 左：計測 3 次元データとデジタル地図とのマッチング結果，中央：時空間画像解析による車載映像とデジタル地図とのマッチング結果，右：デジタル地図への実画像のテクスチャマップ例

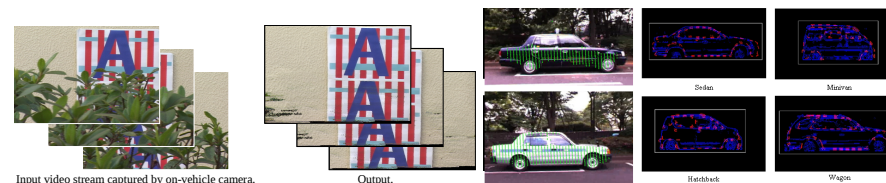


図 2 左：時空間画像解析による障害物除去の例，右：駐車車両検出例

生け垣などを除いたシーンを作り出すことができるようになる。3) では、同様な仮定のもと、建物などのガラス面に写り込んだ鏡面反射光を時空間解析により分離している。また、5) では、常に存在する物体と一時的に存在する物体とを区別する観点から、路上駐車車両の自動認識を行い、また固有窓法を発展させたバイナリフィーチャ法を利用して車種ごとの自動分類を行っている。これらはそれぞれ、幾何学的、光学的、時間的な分離を行った例と言える。実験結果を示した図 2(左) では、複雑な障害物が違和感なく除去されていることが分かる。また、図 2(右) では、駐車車両が正しく検出できていることが確認できる。

2.2 Video-Synth –類似車載映像の検索手法–

ローカルコンピュータに替わって Web 上のクラウドサービスが進化を続けるのと同様に、実世界を計測する仕組みにも変化が生じている。これまでは高価な機材を積んだ特殊な車両がデータを収集していたのに対し、カメラの付いたスマートフォンや自動車の車外周辺モニタ、ドライブレコーダなどが普及したことにより、一般ユーザから収集される都市空間の映像データが Web 上の大規模ストレージに蓄積されるようになってきている。プライバシーに十分に配慮する必要があるが、鮮度が非常に高く、都市部以外からも広く得られると考えられることから、構築した仮想都市空間を更新したり、広い範囲をカバーするなど、これまでのサービスでは難しかったことも実現できる可能性がある。

このようなデータを処理する観点から考えると、精度や密度が十分に保証されないが、その代わり大量にあるのが特徴である。そこでまず、詳細な位置情報の分かっていない都市空間の映像データ群を位置ごとに整理することを考える。7) では、例えば「西新宿」など、

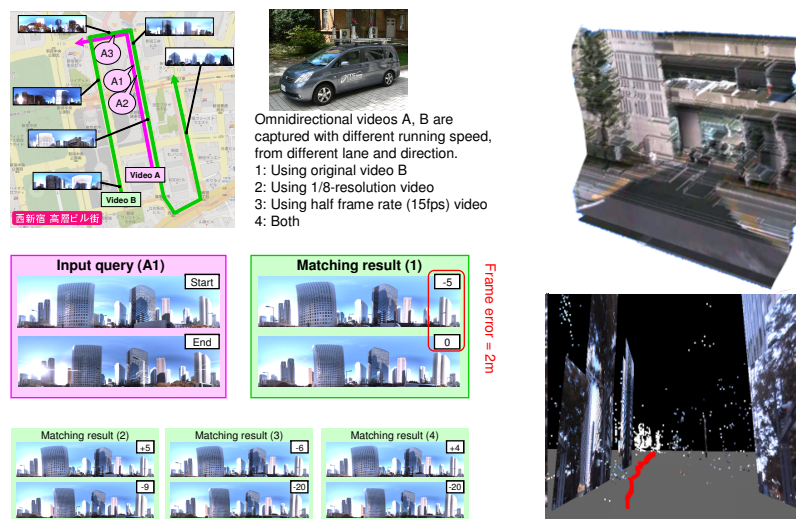


図 3 左：類似映像の検索結果，右：SFM による 3 次元復元および仮想都市の構築例

Web 上で付随するキーワード程度（数 km 程度）の不確定性がある全方位映像データに対し、同じ街路区間が写っていると判断される映像区間をグループ化する方法を試行している。映像の照合には、前述の時系列高画質画像とエピソード平面画像 (EPI) を組み合わせたコンパクトな特徴量として「時空間特徴量」を導入し、これを用いて連続 DP マッチングを行っている。図 3(左) にマッチングの様子を示す。正しく類似映像を検出しマッチングできていることが分かる。

また、ある区間と同じ区間が写ったデータ系列が新たに生じた場合、それを利用して仮想都市空間コンテンツを増強したり高精度化することが考えられる。6) では、単独の全方位映像データから因子分解法により構築した三次元のシーンに対し、別の全方位映像データを追加してバンドル調整をかけ直すことにより、復元精度を向上させることができることを確認した。図 3(右) に 3 次元都市の復元例を示す。

3. 3 次元データの Web 有効利用

3 次元データの Web 利用は、Internet 初期の VRML に始まり、最近の HTML5 の WebGL に至るまで、試みられ続けてきたが、以下に述べる理由より広く普及するに至っていない。

(1) 3 次元コンテンツの作成が容易でないこと

表 3 3D データを用いた Web システムの提案手法

「サービス名」	「集め方・構築方法」	「利用方法・解析方法」
Shape rank	Web 上の 3D + 計測した 3D	3 次元特徴量 + PageRank + 検索
パーソナル Web 3D	計測した 3D + social data (Web の集合知)	自動アノテーション + 視点依存情報提示
画像の自動コメント	Web 上の画像 + 感情辞書	画像類似度を用いた画像への自動コメント
3D motion	我々の手法, キネクト	人の動きの解析など

(2) 3 次元データのサイズが大きいこと

(3) ブラウザで表示するにはプラグイン等が必要なこと

(4) 単純な 3 次元オブジェクトにはコンテンツとしての魅力が欠けること

上記のような課題に対して、(2) と (3) に関しては、既に Web 上の大量データと 3 次元コンテンツを利用して効果的に閲覧可能な可視化手法に関する研究が行われている^{23),24)}。そこで、我々は (1) について、市販のプロジェクタとカメラのみで 3 次元計測が可能な簡易なシステムの開発を進めた結果、一定の成果を納めてきた²⁵⁾。また、(4) に関しては、3 次元コンテンツの検索システム²²⁾ および 3 次元コンテンツへの効率的なデータ付与ならびにユーザにとって意味のある付加価値情報の提示を意図したパーソナル Web3D について研究を進めてきた²⁶⁾。以下では、3 次元コンテンツの検索²²⁾ およびパーソナル Web3D システム²⁶⁾ について紹介し、さらに、応用として 3D 情報を用いた景観ナビゲーションや、さらに巨大な 3 次元モーションデータなど、現在我々が取り組んでいる研究内容について紹介する。

3.1 Shape Rank –類似 3 次元形状検索手法–

本研究では、Web 上にある 3 次元オブジェクトを高精度に検索する手法を提案した。3 次元オブジェクトを検索するためには、3 次元オブジェクトの分類や認識が一般に必要と考えられる。例えば、3 次元オブジェクトの認識に関しては、これまで CAD やモデリングソフトで作成された CG を対象として盛んに行われており、2006 年からはコンテスト (SHREC) が開催され認識精度が競い合われている。しかし、これらの手法は基本的には物体の全周形状が完全に獲得されることを前提としており、実環境において、3 次元スキャナなどにより 1 視点からのみ得られた形状については十分に考えられていない。一方で、実物体を計測した 3 次元形状については SpinImage¹³⁾ や CCHLAC¹⁴⁾ (カラー立体高次局所自己相関) を用いた認識法が提案されているが、基本的には同一の物体を認識する手法のため、Web に不特定多数アップされた 3 次元オブジェクトの検索にそのまま利用することはできない。

近年、高精度な画像検索法として Visual Rank が提案されている¹²⁾。これは、Google の検索エンジンでも採用されている Page Rank¹⁸⁾ の画像への拡張である。Page Rank は、「多くの良質なページからリンクされているページは、良質なページである」という考えに基づいて設計されているが、VisualRank も「良い画像は多くの画像と強く結びつき合っている」とい

う仮説に基づいている。そこで本研究では、Page Rank の 3 次元オブジェクト検索への拡張を目的として Shape Rank を提案した。これは、まず、3 次元オブジェクトをキーワード検索によって検索し、次に 3 次元オブジェクトから形状特徴を抽出し、各 3 次元オブジェクト間の類似度計算を行い、最後に Page Rank と同様の方法でランキングを計算するというものである。3 次元形状の場合、視点依存性などの問題があるため、これを解消する必要がある。

3.1.1 視点依存性を考慮した類似度の計算方法

ShapeRank は、基本的には Page Rank と同様のアプローチで、ハイパーリンクの代わりに各形状間の類似度を定義する。これによって、キーワード検索によって求めた形状ランキングの修正を行う。形状間の類似度がより高いものが上位にランキングされるため、キーワード検索によって集まった形状群の代表的な形状が上位に表示されることになる。隣接行列の類似度は、曲率を表す Shape Index¹⁵⁾ と SIFT を組み合わせて画像間の対応点数を基に計算する。全周形状に対しては、正二十面体の頂点それぞれから、類似度全てを計算することで対応する。その場合、他のオブジェクトと比較する際は、最も類似度が大きい視点を採用することとする。

3.1.2 実計測データを用いたランキング実験

ShapeRank の動作確認として、同一オブジェクト (マルスの彫像) の全周形状と様々な視点から撮影した距離画像群とノイズオブジェクトを用いてランキング計算を行った。特徴量の計算は、各距離画像群に関しては、計測時のカメラの視線方向から計算した (全 14 方向)。3 次元オブジェクトの No. 1-14 が個々の距離画像であり、No. 15 が全周形状であり、No.16-20 が、関連のないノイズオブジェクトである。

提案手法によるランキング結果を図 4 に示す。また、上位にランキングされたオブジェクトを図 5 に示す。図 4 から、石膏像の全周形状がノイズに影響されずに 1 位にランクインしていることが分かる。また、ランキング結果の上位には図 5 に示したとおり石膏像と類似性の高い距離画像がランクインしていることが確認できる。しかし、6 位にアヒルの玩具を計測した全周形状がランクインされた。これは、全周形状の場合最も対応点数の多かった視点を採用するため、無関係のオブジェクト間においても誤接続が発生したことが原因と考えられる。また、各オブジェクト間の類似度を辺重みとする重み付きグラフを図 6 に示す。図 6 を見ると、石膏像の全周形状が石膏像の距離画像と類似性が高くなっていることが確認できる。ノイズとしてあてた No.19 の円柱状の資材にもリンクができていますが、他の距離画像に比べて少ないことが分かる。一方で、一部の石膏像の距離画像と全周形状との間にリンクが出来ていないものがある。これは、正二十面体の頂点全ての視点と、対象距離画

No	Rank
images15	0.1916
images2	0.0763
images6	0.066
images12	0.0661
images1	0.0604
images20	0.0568
images13	0.0544
images9	0.044
images3	0.0431
images7	0.0416
images14	0.0415
images1	0.0406
images5	0.0372
images10	0.0355
images4	0.0336
images18	0.0252
images17	0.0246
images19	0.0225
images8	0.0199
images16	0.0169

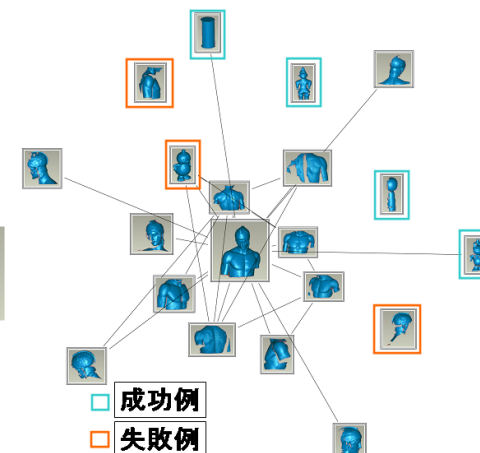
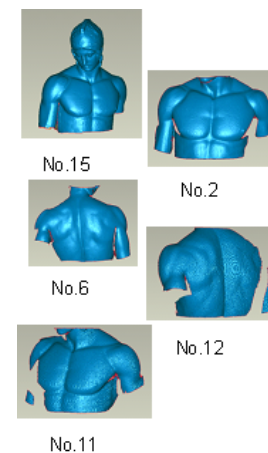


図 4 ランキング結果 図 5 ランキング上位オブジェクト 図 6 類似度を辺重みとする重み付きグラフ

像の視点が離れており、多視点による角度変化の吸収が十分でなかったものと思われる。上記のように幾つかの問題が残されているが、これらの解消が今後の課題である。

3.2 パーソナル Web3D

これまでの主な 3 次元コンテンツに関する研究は Web ページを効果的に提示するための手法が多く、3 次元コンテンツに対する属性情報 (メタデータ) 付与に関する議論は行われてこなかった。このため、3 次元コンテンツにメタデータを付与しようとすると、コンテンツ作成者は 3 次元オブジェクトを作成するだけでなく、3 次元オブジェクトに関するメタデータも作成しなければならず、大きな負担となっていた。一方、Web 上には 3 次元オブジェクトに関する付加価値の高い情報が存在する。そこで、我々はそれらの情報を効率よく収集し、3 次元オブジェクトに容易に付与する手法を提案した。これにより、3 次元コンテンツ作成者は 3 次元オブジェクトに対して低コストで大量のメタデータを作成することができ、コンテンツの価値を手軽に高めることができる。以下では、メタデータの作成支援・表示方法・比較方法それぞれについて紹介する。

3.2.1 3 次元オブジェクトに対するメタデータの作成支援

計測した 3 次元オブジェクトへ Web の情報を用いることで、メタデータ付与を効率的に支援することができる。例えば、3 次元コンテンツ作成者は計測したオブジェクトの名称 (例えば、LUMIX デジタルカメラ) を入力するだけで、オブジェクトの属性となる各種機能 (レンズ、フラッシュ等) の情報を管理するデータベースを自動で作成できる。

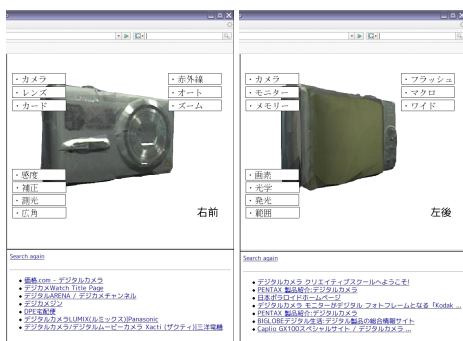


図 7 視点に応じたメタデータの表示例



図 8 メタデータの選択

まず、作成者はオブジェクトの名称を入力する。サーバはその名称に「仕様」「仕様表」などの補間語をキーワードとして追加して Web の検索結果を取得する。次に、ランキング上位のページを取得し、形態素解析によりページを分析し、単語の出現頻度である tf 値を算出し、 tf 値の高い単語を物体の属性名として作成者に推薦する（デジタルカメラの場合には「レンズ」「画素」等）。作成者は推薦されたキーワードから適当なものを選び、3次元オブジェクトの各側面（デジタルカメラの場合は六面）の機能として対応付ける。この対応付け作業により、各側面に対応するテーブルが作成され、対応付けられた機能は各テーブルの属性名となる。さらに、これらの属性名を基に Web から再検索し、返されたウェブページを属性値としてデータベースに格納する。これにより、3次元オブジェクトに対してメタデータが付与される。 tf 値に基づく3次元オブジェクトの属性名の抽出方法は簡単でありながら、有効であることを実験で確認している。

3.2.2 視点に応じたメタデータの効果的表示システム

前記付与されたメタデータを効果的に提示する手法として、視点に応じたメタデータの表示手法を開発した。これにより、3次元オブジェクトを各側面から見た場合では、同じ対象にも関わらずそれに関連する情報が異なっているという、メタデータの視点依存特性を実現することができる。ユーザは3次元オブジェクトを回転させ、ある時点で止めることで、ユーザがその時点において見えている側面を決定できる。具体的には、3次元オブジェクトの中心を始点、ユーザ視点を終点としたベクトルと、各側面の法線ベクトルとのなす角度を計算する。その角度が小さい場合は、その側面が見えているということになる。見える側面

が認識されると、メタデータを保存しているデータベースから、対応する側面に付与されたメタデータをユーザに推薦する。また、一般に1つの属性名に対して、3次元オブジェクトに付与された属性値（収集した Web ページ）の数は膨大であり、その中からユーザに提示するものを選択する必要がある。そこで、ユーザの興味に合ったウェブページを上位にランクし、推薦するシステムとなっている（パーソナライゼーション機能）。

図7に例を示す。ユーザは3次元デジタルカメラ画像を回転させ、右前面で止めると、右側と正面に付与された属性名（レンズ、フラッシュ、広角など）と属性値（関連するウェブページ）がユーザに提示される（左図）。一方で、3次元画像の左後方で止めると、左側と背面に付与されたメタデータが表示される（右図）。さらに、ユーザがデジタルカメラの右前面を選択した場合の表示例を図8に示す。ユーザが1つの属性値（URL）をクリックすると、対応する Web ページが左の領域に表示される。閲覧する度に、Web ページの関連度が再計算され、URL リストが変わりつつ、ユーザの興味と近いページが上位にランク付けされる。実際に被験者によるユーザの閲覧履歴に基づく情報推薦を行ったところ、10回目のアクセス結果において、適合率（妥当な件数/推薦した件数）60%と効果を確認することができた。

3.2.3 複数の3次元オブジェクトの比較

このようなメタデータを用いると、複数の3次元オブジェクトを同時に比較することができる。例えば、側面ごとの比較推薦では、複数の3次元オブジェクトに対してユーザが側面を指定した際に、共通属性と相異属性を特定することや、その側面と他のオブジェクトの側面とのメタデータから類似側面の自動認識が実現できる。図9では、デジタルカメラと携帯電話の3次元オブジェクトが同時に提示されている場合に、ユーザがデジタルカメラの正面と携帯電話の正面を選ぶと、デジタルカメラと携帯電話の正面に付与された共通の属性と各オブジェクトの独自の属性が識別され表示されている例である。

図10は、デジタルカメラの3次元オブジェクトの背面を指定した場合、デジタルカメラの背面と類似する機能を持つビデオカメラの裏面と携帯電話の裏面を自動的に識別することができた画面である。対応側面の自動認識により、ユーザは類似する機能を持つ3次元オブジェクトの側面を容易に比較できる。

3.3 景観を考慮したドライブナビゲーション

ITS や VR などの分野においてナビゲーションシステムの高度な利用方法に関する研究が活発に行われており、最短経路探索や渋滞回避以外にも、個人の興味に応じた観光案内ルート探索なども行われるようになってきている。一方で、ユーザの嗜好や興味に基づいた個人適応化技術の研究が盛んになってきており、それらを基盤とした Web システムにおける



図9 複数の3次元オブジェクトの側面をそれぞれ指定する場合の比較



図10 複数の3次元オブジェクトの対応側面を自動的に認識する場合の比較

情報推薦技術などが実際に利用され始めている．ところが、これまでのシステムでは、Web上の雑多で大量の情報の中から訪れる名所（スポット）を発見するものが多く、局所的なスポットの推薦が中心であり、ルート探索もスポットからスポットへ移動するものがほとんどであり、その移動区間の景観が考慮されることはなかった．ここでいう「景観」とは「風景として遠方から眺めることを目的としたもの」を指す．このようなルート上の景観情報は、人が実際のドライブルートを選択する際の重要な要素の一つである．そこで我々は、Web上のデータから推薦される名所（スポット）などの情報をもとに、単にそのスポットどうしをつなぐことでルートを作成するだけではなく、それらスポット間の移動中の景観も同時に考慮した、借景ルート探索システムを提案した．提案システムは、(1) 任意の地域におけるスポットと景観情報の推薦と、(2) スポットと景観との可視性を考慮したルートの探索・推薦、の2つの特徴を持つ．

3.3.1 景観ナビシステムの概要

提案するシステムは、ドライブに際して、スタート地点とゴール地点、およびドライブしたい地名の3つを入力すれば、Webから収集した観光地情報や景観情報を基に、ユーザとのインタラクティブな操作を通して、ユーザの嗜好にマッチしたドライブルートを推薦するものとなっている．特にドライブ中に車窓から見える景観に注目してルートを探索する方式であり、これは過去の名所を訪問するタイプのルート探索システムとは大きく異なる．

システムの流れを図11を用いて説明する．まず、ユーザは任意の地名を入力する．システムは指定された地名の周辺のスポットと景観をWebから探索して推薦する．次に、ユー

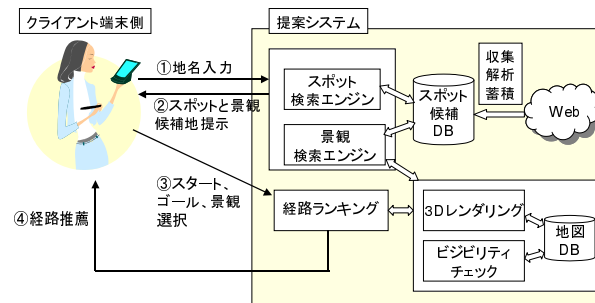


図11 景観ナビシステムの概要

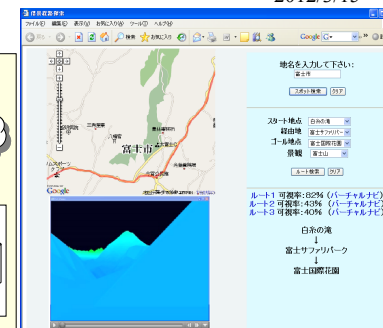


図12 インタフェース．左上に経路を表示し、左下で仮想カメラから見える風景を映像で自動表示．

ザは推薦されたスポットから巡りたいスポットを経由地として選択し、同時に見たい景観を選択する．選択された情報を基に、システムは各スポットを巡る経路と、その経路から景観となる物体がどれだけ見えるか（可視率）を算出し、可視率の高い順に提示する．

クライアントのWebインターフェイス例を図12に示す．左上にある地図が、ルート探索の結果を視覚的に表示したものである．また左下には仮想3次元空間上で経路を通ったときの景観が動画として表示されている．インタフェースの使用法は以下の通りである．

- 地名入力フォームに任意の地名を入力し send ボタンを押すと、セレクトボックスにスポット候補が表示される
- 表示されたスポット候補一覧から経由したいスポットをユーザが選択
- ルート探索ボタンを押すと、ルート探索・可視率計算がスタートし、その結果、ランキングおよび地図と映像が表示される

3.3.2 スポットおよび景観抽出実験

ユーザが地名を入力すると、システムは入力された地名と関連するキーワードを検索キーワード集合として、検索エンジンに問い合わせ、検索結果のWebページをその地名周辺のスポット候補として取得する．関連するキーワードとしては、(1)「観光」、「スポット」、「景観」の3種類の単語と、(2) 検索結果から抽出したスポット名、とした．(2)の場合、再帰的な問合せとなる．この取得した検索結果のWebページから、形態素解析し固有名詞（一般、地域、組織）を抽出し、抽出した単語の tf (term frequency) 値を算出し、 tf 値の閾値の高い順にランキングすることでスポットを取得する．

実験では、日本の都道府県名などの広い地域名を用いて適合率を計測したところ、平均は63%であった．これは、収集した検索結果の上位のページが全国の観光総合案内サイトのページが多いためと考えられる．一方、市町村名など比較的小さい地域名では、適合率の平

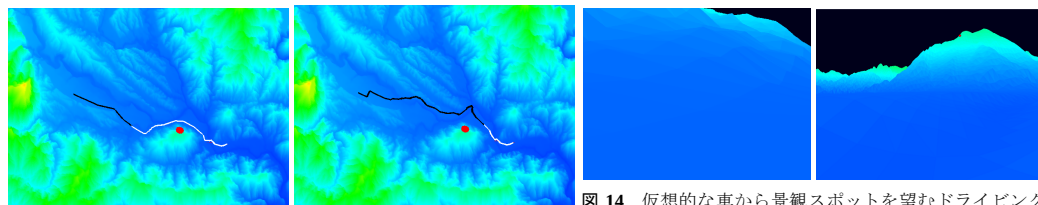


図 13 経路地の異なる 2 つのルートにおける可視性の判定結果。白は目標が見えており、黒は見えていないノード。左が経路 A、右が経路 B。

図 14 仮想的な車から景観スポットを望むドライビングシーンの例。(左：経路 A-斜面に遮られており景観スポットを見ることが出来ない。右：経路 B-見通しがよく景観スポットを正面に見ることが出来る。)

均は 86%であった。これらのことより、日本の市町村名や海外の観光地名に対して、有効な結果が得られることが分かった。また、外国の地名では特にカタカナ表記の場合においても、多くの有効な結果を得られることも確認した。

3.3.3 可視率計算およびルートのランキング実験

埼玉県秩父郡において武甲山を景観スポットに指定し、2 つのルートについて可視率の計算及びランキングの実験を行った。3 次元情報には DEM を利用し、経路を 50m 間隔で正規化した。図 13 は、経路地の異なる 2 つの経路の可視性の判定結果を地図上に表したものである。画像中の丸い点が景観スポットで、経路上の白色は景観が見えている部分、黒色は見えていない部分を示している。継続している可視区間を計算した総可視率を用いると、それぞれ 96.7%(経路 A) と 68.8%(経路 B) となった。経路 A と B は並列して通っている道路であるが、総可視率には約 28% の差がついた。これは、図 14 のように、経路 A が B よりも景観の山の麓を通っているために、景観対象となる点が山の斜面に遮られて見えなくなっているためと考えられる。

3.4 これからの大量 3 次元データ：3 次元モーションデータ

新しい種類の大規模データとして、今後の研究が期待されるものに、動きを含めた 3 次元形状データが挙げられる。既に、3D 映画や 3D TV の規格として、複数視点の動画データや、距離情報を含む動画データの圧縮形式が、MPEG 標準で定められている。

ステレオカメラによる動画撮影は、すでに市販品で可能である。しかし、これはステレオカメラの撮影結果をそのまま提示しているだけであり、より自由な視点での視聴を行うには、距離画像情報など、3 次元形状の解析が必要となる。動的な 3 次元形状の取得は、データ量が非常に大きいこともあり、最近まで高精度かつ安定に行うことが難しかった。しかし、近年、時系列的な連続性の制約を利用して多眼ステレオを効率的に行う研究¹¹⁾や、

TOF カメラや構造化光の利用などのアクティブ計測法^{9),16),17),19)-21)}によって、実用的な計算コストでの形状データの時系列計測が可能になってきた。我々は、格子状パターンを構造化光源としたワンショット計測 (一枚のパターンからの形状計測) による形状計測手法を提案している。例えば、文献²⁰⁾では、グリッド状のパターンと高速度カメラを利用して高速で変化するシーンの形状を、安定でかつ高密度に計測することに成功した。

時系列 3 次元形状のデータ量の大きさは 3 次元コンテンツをネットワーク上で交換、転送する際に、大きな問題となる。また、コンテンツのデータ量を減らすことは、描画コストの削減のためにも重要である。このため、ポリゴンメッシュの圧縮や詳細度 (Level of Detail) の制御を始めとして、多くのデータ圧縮手法が提案されてきた⁸⁾。ところで、3 次元形状は、しばしば点群として表現される。特に、計測によって得られる点群は、スキャンラインや、線状パターンに沿った計測により、曲線の集合として場合が多い。このようなデータでは、曲線の連続性、平滑性から、点群を構成する点の系列には大きな冗長性が現れる。このような冗長性をデータ圧縮に利用する手段として、予測モデルを仮定することが挙げられるが、どのような予測モデルが最適になるかは、データ依存であり、決定することが難しい。我々は、点群系列の予測モデルを多数用意し、切り替えながら利用することで、幅広い種類のデータを効率よく圧縮する手法を提案した¹⁰⁾。図 15 は、本手法で圧縮された点群データが、異なるビットレート (一点を表すためのビット数) において、どのような品質で保存できるか示したものである。

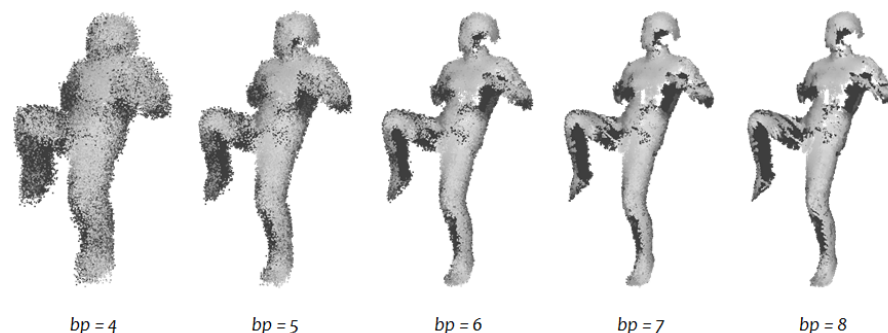


図 15 ビットレートの変化による形状品質の変化。

4. おわりに

本講演では大量データを用いた Web サービスの課題やそれを解決するために我々が取り組んできた事例を紹介した。具体的には、実世界地図データおよび 3 次元データを対象とし、効率的な構築方法や効果的な利用方法をセットで研究開発してきた、これまでの成果を中心に、最近の技術動向についても説明した。この講演が、これからの大量データに基づく Web サービスの発展や理解、さらには大量データが前提となるクラウドコンピューティングやユビキタスコンピューティングなど新しいテーマに対する一助となれば幸いである。

謝辞 本研究の一部は、内閣府・最先端・次世代研究開発支援プログラム (LR030)、総務省戦略的情報通信研究開発制度 (SCOPE) ICT イノベーション創出型研究開発 (101710002) および若手 ICT 研究者育成型研究開発 (102107001) ならびに文部科学省科研費 (21700183, 21200002) の助成を受けて実施されたものである。ここに記して謝意を表す。

参 考 文 献

- 1) L. Tong, S. Ono, M. Kagesawa, K. Ikeuchi, "Matching Range Images of Streets with a Residential Map", IAPR Conference on Machine Vision Applications, pp. 303–306, 2007.5
- 2) 王金戈, 小野晋太郎, 池内克史, 「時系列高画質の提案とそれを用いた車載カメラ画像と建物モデル間の対応付け」, 電子情報通信学会論文誌, vol. J92-D, no. 8, pp. 1197–1207, 2009.8
- 3) T. Oo, H. Kawasaki, Y. Ohsawa, K. Ikeuchi, "Separation of Reflection and Transparency Based on Spatiotemporal Analysis for Outdoor Scene", IPSJ Transactions on Computer Vision and Image Media, Vol. 47, No. SIG9 (CVIM14), 2006.8
- 4) 栗林宏輔, 小野晋太郎, 川崎洋, 池内克史, 「平面群構造を利用した移動カメラ映像からの前景物とテクスチャ面の自動分離」, 画像の認識・理解シンポジウム, 2010.7
- 5) M.Y. Shirmila, S. Ono, M. Kagesawa, K. Ikeuchi, "Fusion of a Camera and a Laser Range Sensor for Vehicle Recognition", 6th IEEE Workshop on Object Tracking and Classification Beyond and in the Visible Spectrum, 2009.6
- 6) 松久亮太, 川崎洋, 小野晋太郎, 阪野貴彦, 池内克史, 「複数系列の車載全方位カメラ画像の対応付けによる広域都市モデル構築」, 画像の認識・理解シンポジウム, 2009.7
- 7) 小野晋太郎, 松久亮太, 川崎洋, 池内克史, 「時空間特徴を用いた車載カメラ映像マッチング」, 画像の認識・理解シンポジウム, 2010.7
- 8) Alliez, P., Cohen-Steiner, D., Devillers, O., Levy, B., and Desbrun, M.: Anisotropic Polygonal Remeshing, ACM SIGGRAPH (2003).
- 9) Canesta, Inc.: CanestaVision EP Development Kit. <http://www.canesta.com/devkit.htm>.

- 10) Daribo, I., Furukawa, R., Sagawa, R., Kawasaki, H., Hiura, S. and Asada, N.: Efficient Rate-Distortion Compression of Dynamic Point Cloud for Grid-Pattern-Based 3D Scanning Systems, 3D Research Journal (2012).
- 11) Furukawa, Y. and Ponce, J.: Dense 3D Motion Capture from Synchronized Video Streams, CVPR (2008).
- 12) Jing, Y. and Baluja, S.: Visual Rank: Applying Page Rank to large-scale image search, IEEE Transaction on Pattern Analysis and Machine Intelligence, pp.1877–1890 (2008).
- 13) Johnson, A.E. and Hebert, M.: Using Spin Images for Efficient Object Recognition in Cluttered 3D Scenes, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, pp.433–449 (1999).
- 14) Kanazaki, A., Harada, T. and Kuniyoshi, Y.: Partial Matching for Real Textured 3D Objects using Color Cubic Higher-order Local Auto-Correlation Features, Eurographics Workshop on 3D Object Retrieval 2009, pp.9–12 (2009).
- 15) Koenderink, J.J. and van Doorn, A.J.: Surface shape and curvature scales, Image and Vision Computing, Vol.10, No.8, pp.557–564 (1992).
- 16) Mesa Imaging AG.: SwissRanger SR-4000. <http://www.swissranger.ch/index.php>.
- 17) Microsoft: Xbox 360 Kinect. <http://www.xbox.com/en-US/kinect>.
- 18) Page, L., Brin, S., Motwani, R. and Winograd, T.: The PageRank Citation Ranking: Bringing Order to the Web., Technical Report 1999-66, Stanford InfoLab (1999). Previous number = SIDL-WP-1999-0120.
- 19) Pan, J., Huang, P.S. and Chiang, F.-P.: Color-coded binary fringe projection technique for 3-D shape measurement, Optical Engineering, Vol.44, No.2, pp.23606–23615 (2005).
- 20) Sagawa, R., Kawasaki, H., Furukawa, R. and Kiyota, S.: Dense One-shot 3D Reconstruction by Detecting Continuous Regions with Parallel Line Projection, Proc. 13th IEEE International Conference on Computer Vision (ICCV 2011), pp.1911–1918 (2011).
- 21) Weise, T., Leibe, B. and Gool, L.V.: Fast 3D Scanning with Automatic Motion Compensation, Proc. IEEE Conference on Computer Vision and Pattern Recognition, pp.1–8 (2007).
- 22) 石川雅浩, 鈴木崇之, 下野博, 古川亮, 河合由起子, 川崎洋: 形状特徴に基づく web 上の 3 次元オブジェクトの効率的検索手法の提案, インタラクシオン 2010 (2010).
- 23) 灘本明代, 矢部武志, 四方正輝, 田中克己: 3 次元 CG コンテンツとその属性情報の自律的呈示方式, 情報処理学会誌, Vol.43, No.SIG2(TOD13), pp.203–215 (2002).
- 24) 小谷 彬, 小山 聡, 田中克己: 複数 Web コンテンツの多面的閲覧のための空間インタフェース, 第 16 回データ工学ワークショップ (DEWS2005) 論文集, pp.6C–i7 (2005).
- 25) 川崎 洋, 大沢 裕, 古川 亮, 中村泰明: 空間コード化法を用いた未校正ステレオシステムによる密な 3 次元形状復元, 情報処理学会論文誌 CVIM, Vol.47, No.10, pp. 59–71 (2006).
- 26) 鈴木崇之, 下野博, 河合由起子, 川崎洋, 石川雅浩, 張建偉: 協調型 Web3D コンテンツ管理システムの実装, インタラクシオン 2010 (2010).