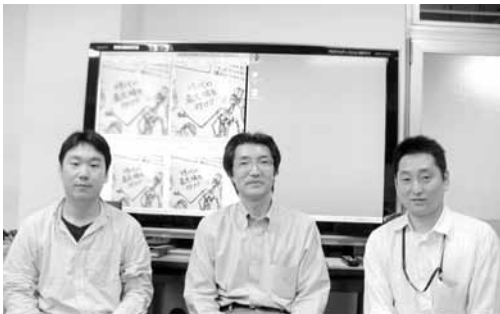




進歩し続ける画像処理

奥富・田中 研究室～機械制御システム専攻



奥富 正敏 教授 (中央)

田中 正行 准教授 (左)

杉本 茂樹 研究員 (右)

現代の先進技術の一つとして、デジタル画像に関する研究が盛んに行われており、テレビやデジタルカメラなどの映像機器の性能は著しく向上している。また、カメラを用いて取得できる画像情報は、人間の視覚情報と同様に大きな情報量を持つため、さまざまな方法で活用されている。

奥富・田中研究室はそのような技術の研究の一端を担っている。この記事では画像の品質を向上させるため、また、画像情報を用いて自動車の運転支援をするため、奥富・田中研究室が行っている研究について紹介する。



デジタル画像の利用法

現代では、さまざまなものがデジタル化されている。画像もその一つだ。一昔前では写真を撮影するとき、フィルムを使いアナログ画像として記録していた。一方、現代ではデジタルカメラやスキャナなどを使い、デジタル画像として記録している。アナログ画像と異なり、デジタル画像はコンピュータ上で編集を施すことで、多種多様に変化させることができる。奥富・田中研究室では、このデジタル画像の画像処理について、多くの研究を行っている。

デジタル画像はアナログ画像に比べて粗い。これは、アナログ画像が色の濃淡や明るさなどの連続的な変化によって成り立っているのに対し、デジタル画像が単一の色を持つ小さなブロックの集まりで構成されているからである。このブロックのことを画素と呼び、また、画像を構成する画素の数を解像度という。この解像度が高いほど、デジタル画像でより細かに表現することができる。一方、解像度の低い画像では何が示されているのかを判別できない場合がある。そこで、解像度の低い画像に何らかの処理を施し、解像度を高くする方法が世界中で研究されている。

ある程度解像度の高いデジタル画像は、人間が視覚によって得られる画像と同様に、多くの情報を持っている。例えば、人間があるデジタル画像に写っている個人の顔を認識できるとする。これは画像がその顔を判別するのに十分な情報を持っているということだ。この、解像度の高いデジタル画像が人間の見ている画像に近い情報量を持っていることを応用すると、さまざまなことが可能になる。例えば、人間は二つの目で見た映像の微妙な違いから、目の前の世界を立体的に把握している。これと同様に、少し離れた二つのカメラで撮影した、解像度の高い同じ風景の画像2枚を用いれば、撮影された風景の立体的な配置を推測することも可能だ。

奥富・田中研究室では、複雑な画像処理を用いてこれらを実現する方法を研究している。その他にも、月面の上空写真から地形を三次元的に再現する方法や、歩行ロボットの床面情報の取得方法、デジタルカメラの手振れを補正する方法など、研究は多岐にわたる。これから、先に紹介した、画像の解像度を高める研究と、自動車の積載したカメラによる運転補助の研究を紹介しよう。

超解像処理

近年、テレビやパソコンなどのディスプレイの高解像度化により、画像の解像度をより高めることが必要とされている。監視カメラで撮影された人の顔や、車のナンバープレートなどの認識性能を上げるためにも、解像度の向上に対する要求は日に日に高まるばかりである。

しかし、カメラの光学センサの性能を向上させて解像度の高い画像を得るには、技術的、経済的な限界がある。そこで、低解像度の画像から高解像度の画像を生成する処理が必要となった。奥富・田中研究室では、画像の解像度を高める方法の一つである、超解像処理を研究している(写真1)。ここではまず、いくつかある超解像処理に共通する基本的な考え方を述べる。その後、奥富・田中研究室が提案する画像処理の高速化アルゴリズムを簡単に紹介しよう。

1枚のデジタル画像はその画素の数しか情報を持っていないので、それよりも高い解像度の画像を生成するためには情報が足りない。そこで、対象物を撮影した画像を複数用いて情報を補完することによって、より高い解像度の画像を生成する。必要となる複数の画像は対象をある程度同じ角度から記録した動画から得ればよい。



超解像処理前



超解像処理後

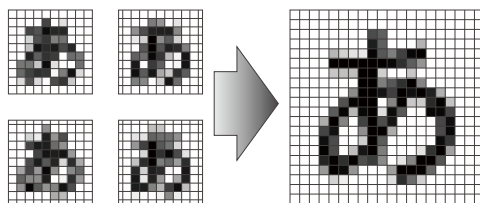
写真1 超解像処理を施した写真

図1に、4枚の同じ解像度の画像を基に、1枚の高解像度の画像を生成する簡単な例を示す。4枚の元画像の画素がちょうど半画素ずつ互いにずれていた場合、この4枚の画像を単純に重ね合わせるだけで、 2×2 倍の解像度の画像を得ることができる。このように、複数の画像を統合して高解像度画像を生成する処理を再構成処理という。

元画像がちょうど半画素ずつずれていれば、再構成処理は簡単にできる。しかし、多くの場合、適当に撮影された画像は正確に半画素ずつずれているわけではない。それらは平行移動、回転、スケール変化などさまざまなずれ方をしている。そのため、再構成処理をする前に、複数の元画像の位置を合わせる処理が必要となる。

位置合わせ処理では、ある基準となる画像と、変化させたい1枚の入力画像を用意する。これら2枚の画像の差が小さくなるように、入力画像を変形させるパラメータを更新していくのが、位置合わせ処理の基本的な方法である。基準画像と入力画像の差は、画素値(画素の持っている色情報)の類似度を用いて評価される。この類似度を最大化することで位置合わせ処理となる。このようにして、位置合わせ処理は、二つの画像と変形パラメータを考慮した評価関数の最小化問題として定式化され、これを解くことによって画像の位置を合わせることができるのだ。

再構成処理の詳しい説明の前に、カメラモデルについて説明しておこう。画像は実シーンをカメラで撮影することによって得られる。このとき、カメラは実シーンから発せられた光をレンズに通して、光学センサで読み取り画像を生成して



再構成処理により、半画素ずつずらして撮影した4枚の10×10の画像から、20×20の画像をつくる

図1 半画素だけずれた画像同士の再構成処理

いる。この過程で、画像は実シーンに対して正確に生成されるわけではない。通常デジタルカメラで撮影する際、各画素ごとに光学センサが三原色のRGBを読み取り、画素値にしている。しかし、センサ一つは三原色の一色分しか感知できず、三色分のセンサを同一箇所に配置することもできない。そのため、画素の中心からずれた位置にある三つのセンサの情報から、それぞれ色情報を補完して、その値を各画素の値として用いることになる。こうして撮影画像と実シーンは微妙に異なってしまう。また、レンズを通った光は若干ぼやけており、カメラの構造による誤差を含んでいることも、実シーンと画像との誤差を生む。カメラモデルとはこのような、あるカメラで撮った画像と実シーンの間の対応関係を表すものである。

再構成処理では、生成する高解像度画像の画素値はすべて未知であるため、近似を用いずに計算して決定することが非常に難しい。そこで、次のような方法をとる(図2)。まず初めに、目標となる高解像度画像の画素値を仮定した後、観測画像を生成する。観測画像とは、前述の位置合わせ処理を施した複数の低解像度画像を重ねた画像のことであり、その中では画素値を持つ点が散在している。次に、比較用に仮定した、高解像度画像にする予定の画像にカメラモデルを適用させ、観測画像の各点と対応する点の画素値を計算する。こうして求めた仮の画像の画素値と、観測画像の画素値の差が小さくなるように、仮の画像を修正する。この修正された画像を基にして、修正を何度も繰り返すことにより、実シーンに近い高解像度画像を得ることができる。

以上に説明してきた超解像処理であるが、これには計算に長い時間がかかるという欠点がある。

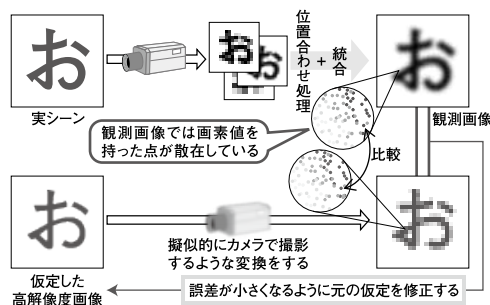


図2 超解像処理の流れ

特に再構成処理では、観測画像の全ての画素に関して高解像度画像にカメラモデルを適用させる計算を何度も繰り返すため、計算量が膨大である。そこで、奥富・田中研究室が提案している、再構成処理を高速化するアルゴリズムを紹介しよう。

奥富・田中研究室が提案しているアルゴリズムでは、高解像度画像にカメラモデルを適用させて画素値を算出する計算の回数を減らすことで、全体の計算時間を短縮する。従来では、観測画像の画素値を持つ全ての点について、その画素値を算出していた。このアルゴリズムでは、処理の精度を維持しつつ、この点の数を少なくする。具体的にはまず、観測画像において高解像度画像の1画素に対応する領域を考える。この領域に存在する全ての点の代わりに、それらの画素値を平均化した画素値を持つ一つの点で置き換える。このとき、単に一つの点で置き換えただけでは、置き換えられた高解像度画像の点が参照した観測画像の点の数を考慮しておらず、その点の情報がいかに正確かわからなくなる。そこで、点の数を考慮した重みを置き換えられた点に与えることで、この区別ができるような工夫をしている。これにより、観測画像の全画素の数だけ行わなければならない処理の回数を、高解像度画像の画素の数だけに減らすことができる。このアルゴリズムは元の処理との誤差が非常に小さく、超解像処理としての質はほとんど劣化しない。それにもかかわらず処理速度は従来に比べて約1.3～5.0倍となる。奥富・田中研究室では、このようなアルゴリズムによって、画像処理を高速化しようとしている。

この他に、奥富・田中研究室では超解像処理を効果的に行う研究をしている。通常、画素値はRGB系で扱われると述べたが、奥富・田中研究室では明るさを表す輝度Yと、色差を表すCb,Crで色を表すYCbCr系に変換して扱っている。人間の目は色彩の変化よりも明るさの変化に敏感であるため、このように色を輝度とそれ以外に分けて処理することにより、人間の目で綺麗に見える画像を生成することができるのだ。

ここで紹介した高速化アルゴリズムや処理を行う上での色の工夫は、奥富・田中研究室における超解像処理の研究成果のごく一部である。ここでは紹介できなかったが、この他にもさまざまな方法を用いて超解像処理の向上を目指している。



カメラを使った運転支援

近年、運転手の負担を軽減することを目標として、自動車のインテリジェンス化を進める研究が行われている。例えば、走行中の車両の位置を道路上の白線の間に維持させる機能や、細かい操縦を要求される駐車支援や、衝突防止に利用できる自動ブレーキ機能、視界の悪い環境で歩行者を検出する機能などがある。また、車両内の運転手のわき見や居眠りを検知し、運転手にシートベルトで刺激を与えるという機能のように、さまざまな機能が考案されている。

これらの機能の中には、車両より前方にある障害物の位置やそこまでの距離の情報をを用いるものがある。具体的には、前走車との距離を一定以上に保つ機能や、前走車の急ブレーキに反応して自動的にブレーキをかける機能などである。位置情報の取得には、主にレーザレーダなどが用いられている。しかし近年では、より安価で空間的な解像度が高く、色や模様といった情報も得られるという特性を持つ、カメラを用いた技術も盛んに研究されている。例として、車両前方の左右にカメラをおき、そこから得られる画像を処理することで位置情報を取得する方法が挙げられる。これらはステレオカメラと呼ばれ、カメラを一つしか使わない場合と違い、対象との距離を正確に把握することができる。

奥富・田中研究室では、ステレオカメラで得られた二つの画像を位置合わせするため、道路面を

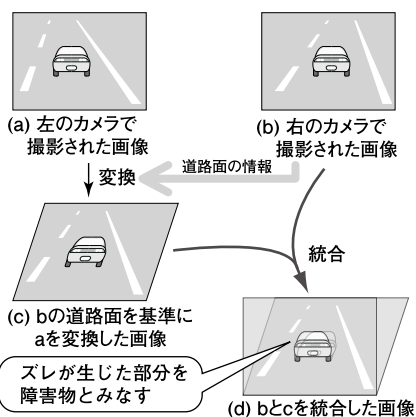


図3 ステレオカメラで得られる三次元情報

基準とした変換を行っている。基本的に道路面は平面と見なすことができるため、画像の位置合わせをする上で基準にするのに最適である。撮影されたいずれかの画像の道路面を基準とし、それに合うようにもう一方の画像の道路面を変換する(図3)。ここでも超解像処理で用いた位置合わせの技術を利用している。また、道路面を基準として位置を合わせた画像を重ねると、道路領域の像のみが一致し、道路上の障害物の像にずれが生じる。この重ねた画像を利用すると、画像上から道路領域でない部分、すなわち自車両にとっての危険領域を、画像認識のみで検出できる。

しかし、こうして検出した危険領域の危険性を正しく判断するには、画像上で示された危険領域を空間的な位置に変換し、自車両との相対的な位置を把握することが必要である。そこで、仮想投影面画像を利用する(図4)。この画像は自車両前方の道路面を真上から垂直に見下ろした画像であり、この画像の情報はそのまま空間的な対象の位置や変化を表す。道路面の法線ベクトル、さらにカメラと道路面間の距離から、元の画像と仮想投影面画像の関係を表す行列が求められ、これからカメラと道路面との姿勢の関係がわかる。この関係を利用すると、仮想投影面画像を生成することができる。

仮想投影面画像を用いた危険領域の検出は、この画像における、カメラ間の中点から非道路領域までの距離の変化に基づいて行う。カメラ間の中点から道路領域の端点までの長さが非道路領域までの距離である。この距離と、画像が一定時間ごとに撮影されることを用いると、自車両に対する非道路領域の相対速度を求めることができる。



図4 仮想投影面画像

求められた相対速度が負である場合、それは自車両に相対的に接近していることを表す。よって、対向車や静止した物体など、自車両にとって危険な障害物が存在している可能性が高く、危険領域であると見なせる。相対速度が0の場合は自車両と同じ速度で走る車両など、正の場合は自車両よりも速く走っている車両などと考えられる。

こうして検出された情報は、視覚誘導や安全走行のための運転支援といった機能に利用できる。

これまで述べてきた危険領域検出の方法は、自車両から一定距離の位置を一括して計測する、効率を重視した方法であった。しかし、この検出の仕方では性質上、対象が道路面に接していない場合や、あるいは道路面の形状が平面と大きく異なる場合などでは、障害物までの距離が正しく計測されない可能性があった。このため、障害物を直に検出することで、検出の安定性や精度を重視する方法が考えられた。危険領域検出に対して障害物検出と呼ばれるこの技術を紹介しよう。

障害物を検出する際、走行上衝突の危険がある前走車やガードレールなどは検出し、一方で衝突の危険がない信号機や歩道橋、路面上に描かれた標識などは検出しないようにする必要がある。そこで、自車両と衝突し得る高さに存在するもののみを障害物として区別する方法を考える。

まず、物体が存在する高さの範囲を求めるには、仮想投影面画像を生成する際に用いたような変換を施すことで、フロントガラスと同じ高さにあるカメラを基準とした座標系から、道路面を基準とした座標系に変換して計測を行う。これにより、道路面上の計測点の高さを0として路面から一定の高さの空間を容易に測定することができる。

計測は道路面を基準とした座標系において、障害物のあると思われる高さの範囲を指定し、その中で物体の存在する領域に点をつけて行われる。障害物となる物体はある程度の大きさを持つため、計測点は空間的な集団を形成する。障害物を検出するためには、その集団を検出すれば良い。

次に、クラスタリング処理を行う。クラスタリングとは、計測した点を同じ物体を示すものと

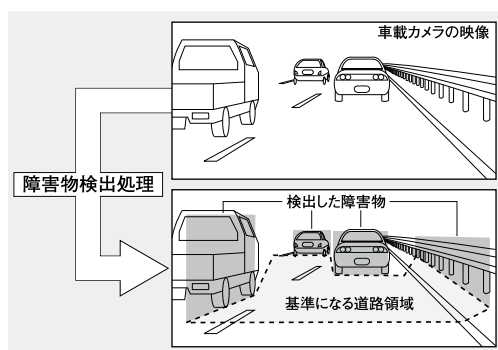


図5 障害物検出結果

にまとめて区分けするものである。ここでは、隣り合った画素同士の類似度が高かった場合にその画素同士をクラスタリングするという、最短距離法と呼ばれる方法を用いている。これによって計測点の集団を生成することができる。

ステレオカメラによる計測では、遠方になるにつれて計測点の密度がより疎になってしまう。よって、クラスタリングの結合範囲を自車両からの距離によらずに一定にした場合、遠方の障害物が検出されなかったり、近接する物体が不必要に結合したりしてしまうことが考えられる。これを防ぐため、近くと遠くでデータの疎密を考慮した補正をかけ、より正確なクラスタリングを行う。これらの結果からノイズを取り除くことで一連の処理が完了する。

実際に障害物検出を行った結果が図5である。障害物検出はステレオ画像中の位置、すなわち実空間中の位置を求めることができるため、自動車の制御や他の処理に利用しやすいと考えられる。

今後の車両前方認識の研究においては、奥富先生は二つの方向性があると考えている。一つは、センサとカメラを併用することで認識処理の性能を上げ、検出率の向上、誤検出率の低減をするといったセンサフュージョンを活用するという方向性である。またもう一つは、自動車や歩行者といった障害物の種類や道路標識の認識などを加えることでシーンのコンテキストを考慮するという方向性である。奥富先生らの今後の研究によって、これらの二方向にさらなる発展が見られるだろう。

今回の取材で伺った画像処理の話は、私たちに身近で大変興味深いものでした。度重なる取材に

親切に応じて下さった奥富先生、田中先生、杉本先生にお礼申し上げます。(山梨 純平)