

車載カメラ画像からの視点変更画像生成に関する検討

石川 諭[†] 小林 祐一[†] 金子 透[†] 山下 淳[‡] 石原 進[†]

[†] 静岡大学大学院工学研究科 〒432-8561 静岡県浜松市中区城北 3-5-1

[‡] 東京大学大学院工学系研究科 〒113-8656 東京都文京区本郷 7-3-1

E-mail: †{f0330003,tykobay, tmtkane } @ipc.shizuoka.ac.jp,

‡yamashita@robot.t.u-tokyo.ac.jp, †ishihara@ishilab.net

あらまし ITS において交通状況の認識・伝達は重要な分野の一つである。車々間通信のアドホックネットワークによって車載カメラ画像をやり取りすることはインフラに非依存でリアルタイムな交通状況の認識が可能になると期待されている。しかし通信状況によっては求められている場所の画像を得ることが出来ない可能性があり、近傍で撮影された画像から求められる場所の画像を生成する方法が求められる。そこで本研究ではこの問題の解決として1枚の車載カメラ画像から仮想視点画像を自動生成する方法を提案する。

キーワード 画像処理, ITS, 車載カメラ, 透視投影変換

1. 序論

1.1. 背景

情報通信技術を利用し道路交通の問題を解決するものとしてITS(Intelligent Transport Systems)がある。ITS において交通状況の認識・伝達は重要な分野の一つであり、その代表的なものに VICS(Vehicle Information and Communication Systems)や ITS スポットサービスがある。

VICS は測定地点が多く広範囲で利用できる一方で蓄積した交通状況のデータを扱うため即時性に欠けるといった問題が存在する。これに対して ITS スポットサービスは、主要道路に設置された定点カメラの映像を車両のカーナビゲーションシステムに表示し、運転者が現在の交通状況を直感的に認識できるもので即時性に優れる。しかし定点カメラの設置にはコストがかかり、測定地点が少なく利用可能範囲が狭いのが問題である。

これらの問題に対して車々間通信によるアドホックネットワークの提案がされている[1]。現在ドライブレコーダーなどの車載カメラを搭載する車両が多数存在する。車載カメラ画像を車々間で通信しカーナビゲーションシステムに表示することで、ITS スポットサービスの利点である直感的な認識のしやすさを保ちつつ測定地点を多くし VICS のように広範囲で利用できる。しかし車々間通信の通信状況は必ずしも万全ではなく、通信の遅延や断絶が発生する。また画像が欲しい地点に必ずしも車両があるとは限らない。そのような状況の際には交通状況を知りたい場所の画像が表示できない事態が発生する。

この問題点に対してカメラが無い場所の付近に車載カメラを搭載した車両がいる場合、その車載カメラ画像を利用し仮想視点の画像の提示ができれば、交通状況の認識がより広い範囲で利用できる。仮想

視点における画像生成手法として Horry らの TIP[2]がある。本手法は1枚の画像から仮想視点画像を生成する手法であるが、画像中の対象物を必要に応じて手動で抽出処理をしている。

1.2. 研究目的

車々間通信での問題点の解決を試みるため、本研究では一枚の車載カメラの画像から前方方向に移動した仮想視点画像の生成方法を提案する。図1のように車載カメラが無い場所に対して、そこより後方に車載カメラを搭載した車両が存在する問題を想定する。後方にある車両の車載カメラの画像から前方方向に視点を移動した画像を自動生成し、その場所の交通状況を認識可能にする画像提示を行う。1枚の画像のみを用いるのはデータ量が動画や複数画像に比べて少なく、即時性を損なわないことが期待されるためである。

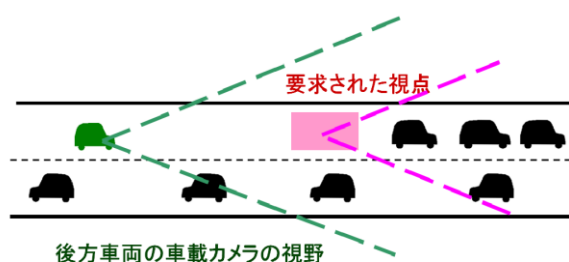


図1 問題設定で想定される道路状況

2. 前方方向視点画像の生成

2.1. 処理手順

取得画像より白線と車両を検出し、画像中の道路領域と車両領域、および消失点を求める。次に求めた消失点を利用して画像を変換する。最後に取得画像の道路領域と車両領域と変換した画像を合成する。

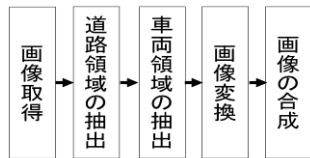


図2 処理手順

2.2. 白線と消失点の検出

取得画像に対して白線検出を行う。前処理として取得画像のエッジ検出を行う。次にエッジ検出画像にハフ変換を行い直線を検出する。角度が正と負の直線のうち、最下部の右から画面幅4分の1の位置の点との距離が最も短く、垂直でない直線を1本ずつ選び、これらを白線とする。この位置を選ぶ理由は右側の白線について、中央線ではなく反対側車線の白線を検出するためである。また2本の白線の交点を消失点とする(図2)。最後に図3のように白線と画面下辺との交点2つと消失点を結ぶ三角形を道路領域とする。この道路領域は2.4節で述べる変換を行わない。

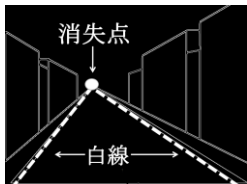


図2 白線と消失点の検出



図3 道路領域

2.3. 車両領域の検出

車両検出には HOG(Histogram of Oriented Gradients) 特徴量[3] を用いる。HOG 特徴量は局所領域の輝度勾配方向をヒストグラム化した特徴量で、人や車両などの物体検出に有効な特徴量である。HOG 特徴量に関して予め車両とそれ以外について RAB(Real AdaBoost)[4] で識別器を構成する。RAB は特徴量の確率密度分布を用いた学習方法である。車両画像とそれ以外のサンプルを用意し、ある局所領域の HOG 特徴量の出現確率をそれぞれ求め、局所領域とその HOG 特徴量あからなる識別器とする。

まず取得画像に対して検出ウインドウ内の HOG 特徴量を得る。その HOG 特徴量から、RAB により車両かどうかを判別を行い、車両と判別した場所を記録する。この処理を検出ウインドウの場所や倍率を変えて繰り返す。車両と判別された検出ウインドウを MeanShift 法[5]でクラスタリングし、クラスタリングされた数が閾値以上なら車両とする。図4はこの手法で車両検出した模式図である。車両として検出された場所と大きさはクラスタリングされたウインドウの位置と倍率の平均である。この場所と大きさを車両領域とし図5のマスク画像を生成する。



図4 車両検出の例

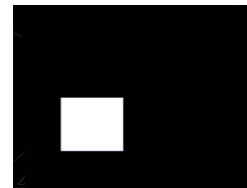


図5 車両領域

2.4. 透視投影変換

1枚の画像から仮想視点画像を自動生成するため、本研究では TIP を参考に透視投影面に投影されているものは左右2枚の壁に描かれたものであるという仮定をおく(図6)。ただし信号機や電線などの道路上部にある物体については2.5節で別途考慮する。

まず座標系の原点は透視投影面上の消失点とする。透視投影面を xy 平面とし、光軸と平行になるように z 軸をとる。次にこの座標系において視点と消失点を通る直線と y 軸に平行に壁が左右に存在し、その壁に描かれた建物などが透視投影面に投影されると仮定する。そして透視投影面と視点が視点と消失点を通る直線と同じ方向へ平行移動した際に推定される投影画像を変換した画像とする。

対応する画素について、2.1節で求めた消失点を原点としたとき取得画像と変換した画像のそれぞれの座標を (u, v) , (u', v') とする。また、カメラの像距離を f 、消失点から壁までの距離を D_i とする。なお i は1が消失点より右側、2が左側を意味する。視点と消失点を通る直線に沿いながら進み、視点と消失点を結ぶ直線方向の移動量を S とすると (u, v) と (u', v') の対応関係は

$$u = u' / \left(1 - \frac{Su'}{fD_i} \right), v = v' / \left(1 - \frac{Su'}{fD_i} \right), i = 1, 2 \quad (1)$$

となる。式(1)の S の値を図1の交通状況を知りたい場所と後方車両の距離に合わせて変更することで、目標とする画像に変換できる。

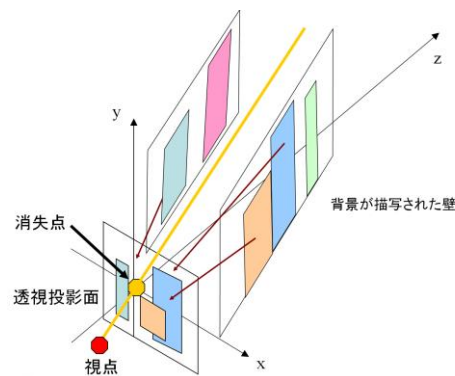


図6 透視投影変換での仮定

2.5. 道路上方の物体

本研究の目的は画像による交通状況の認識である。画像中の信号機をはじめとする交通標識は交通状況

の認識において重要な情報である。交通標識に対して 2.4 節の変換を行うと、画像が不自然に歪む場合がある。その対策として交通標識を 2.3 節の車両のように検出を行い、別途変換を行えばよい。本研究では交通標識の代表的なものとして信号機を採り上げ、信号機が道路上方に在る物体であることに着目して、次の 2 種類の方法を提案する。

一つ目の方法は 2.4 節の 2 枚の壁のモデルに天井を加える方法である。画像中の左右の白線を壁と床境界線とし、壁の高さを一定と考えると TIP では天井の高さは求まる。天井モデルにおける 2.4 節の (u, v) 、 (u', v') の対応関係は式(1)を参考に

$$u = u' / \left(1 - \frac{Sv'}{fC} \right), v = v' / \left(1 - \frac{Sv'}{fC} \right) \quad (2)$$

となる。C は消失点から天井までの距離である。白線と D_1 と C の関係を図 7 に示す。一点鎖線が透視投影面、太い点線が白線を示している。

もう一つの方法は信号機のみを抽出し、式(2)の変換を行った後に合成をする方法である。

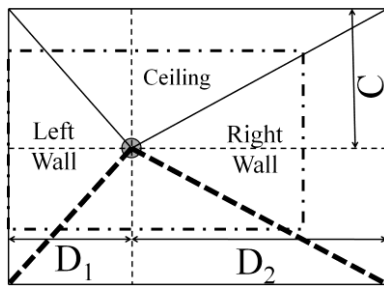


図 7 壁や天井と透視投影面の関係

2.6. 画像合成

道路領域を図 6 での xz 平面に近似するために 2.2, 2.3 節で求めた道路領域と車両領域を利用する。変換した画像に道路領域と車両領域をマスク画像として領域部分に元画像を合成し、道路領域が歪まないようにする。また図 8 のように合成した画像中に同一の車両が重なることを避けるために道路領域外の車両領域を除去する。除去された部分には Navier-stokes 方程式をベースとした手法[6]により補間を行う。道路領域外の車両を除去し補間した画像が図 9 である。



図 8 除去無しの場合



図 9 除去有りの場合

3. 実験

3.1. 実験条件

車載カメラで撮影した 2304×1296 pixels の画像

を使用した。車両検出に用いる RAB による学習は予め行った。またカメラの像距離は $f=840$ と推定した。

3.2. 実験結果

2.1 節～2.4 節及び 2.6 節の手法での図 10 の取得画像に対する実験結果を図 11～13 に示す。図 11, 12 では取得画像に対する白線と車両の検出結果である。図 13 は最終的な生成結果である。平均処理時間は OS: Window8, CPU: Intel Core i7-3770, Memory: 32GB の計算機を用いて 5.7×10^2 ms であり、即時性を失わずに仮想視点画像の生成を達成した。

図 14 は図 6 における z 軸方向への移動量すなわち式(1)の S の値が図 13 よりも大きい変換結果である。S の値が大きいほど道路左右の建物が画面手前に移動しているが、道路との境界線は常に一定の角度になっていることが分かる。問題点として x 軸方向の奥行きが無いため、左右の建物が上下に開く様に歪んでいるのが分かる。これには建物の x 軸方向の大きさが推定できれば改善できると思われる。

次に信号機の存在する図 15 の取得画像に対して 2.6 節で示した二つの手法についての変換結果が図 16, 図 17 である。図 16 では天井部分での信号機や道路をまたぐ電線など水平な物体が違和感が少なく変換されている。その一方で天井と壁との境界をまたぐ水平な物体は境界部分で折れ曲がるという欠点もある。図 17 ではそうした壁と天井の境界での折れ曲がりはない。しかし図 14 と同様に消失点上部で電線が折れ曲がり、信号機を抽出したために信号機が空中に浮くようになっている。



図 10 取得画像



図 11 白線検出

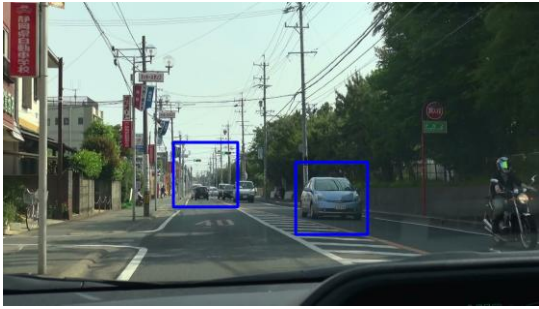


図 12 車両検出



図 17 信号機に対して式(2)を用いた変換結果



図 13 生成画像(S=300)



図 14 生成画像(S=1500)



図 15 信号機のある取得画像



図 16 天井モデルでの変換結果

4. 結論

本研究では一枚の車載カメラ画像からの透視投影を用いて即時性を考慮した前方方向視点画像の生成を達成した。本研究の利点は1枚の市街地の車載カメラ画像から仮想視点画像を自動生成を行ったことである。また今回は本研究の問題点の解決に対して2種類の解決策の提案を行った。今後の課題としては即時性を考慮しつつ2枚の車載カメラ画像を統合することでより広い範囲での仮想視点画像の生成を行うことがあげられる。

謝 辞

本研究は日本学術振興会科学研究費補助金基盤研究 B(23300024)の援助を受けた。

文 献

- [1] 山中麻理子, 石原進: “VANET における Geocast による要求を想定した位置依存情報の複製配布方”, 情報処理学会論文誌, Vol.50, No.1, pp.31-41 (Jan. 2009)
- [2] Y. Horry, K. Anjyo, and K. Arai: “Tour Into the Picture: Using a Spidery Mesh Interface to Make Animation from a Single Image”, SIGGRAPH, Proceedings of the 24th Annual Conference on Computer Graphics, pp.225–232 (Aug. 1997)
- [3] N. Dalal and B. Triggs: “Histograms of oriented gradients for humandetection”, IEEE Conference on Computer Vision and Pattern Recognition, pp.886–893 (Jun. 2005)
- [4] R.E. Shapire and Y. Ahonem: “Improved boosting algorithms using confidence-rated predictions”, Machine Learning, Vol.37, No.3, pp.297–336 (Dec. 1999)
- [5] D. Comanicu and P. Meer: “Mean Shift: A Robust Approach toward Feature Space Analysis”, IEEE, Transactions on Pattern Analysis and Machine Intelligence, Vol.24, No. 5, pp.603–619 (May 2002)
- [6] M. Bertalmio, A. L Bertozzi and G. Sapiro: “Navi-er-stokes, Fluid Dynamics, and Image and Video Inpainting”, IEEE, Computer Society Conference on Computer Vision and Pattern Recognition, Vol.1, pp.355–362 (Dec. 2001)