

時系列情報を用いた車載カメラ画像からの道路標識の視認性推定

道満 恵介[†] 出口 大輔[†] 高橋 友和^{††} 目加田慶人^{†††} 井手 一郎[†]
村瀬 洋[†] 玉津 幸政^{†††}

[†] 名古屋大学 大学院情報科学研究科 〒 464-8601 愛知県名古屋市千種区不老町

^{††} 岐阜聖徳学園大学 経済情報学部 〒 500-8288 岐阜県岐阜市中鶉 1-38

^{†††} 中京大学 情報理工学部 〒 470-0393 愛知県豊田市海津町床立 101

^{††††} 株式会社デンソー 〒 448-8661 愛知県刈谷市昭和町 1-1

E-mail: †{kdomain,ddeguchi,ttakahashi,mekada,ide,murase}@murase.m.is.nagoya-u.ac.jp

あらまし 本稿では、不快感のない自動車の運転支援システムの実現を目的とし、時々刻々と変化する走行環境において道路標識の視認性を高精度に推定する手法を提案する。近年、様々な運転支援システムが自動車に搭載されるようになってきている。しかし、過度の情報提供はドライバーの注意散漫や苛立ちによる不快感の原因となるため、適切な情報の取捨選択が必要である。そこで本研究では、道路交通において重要な道路標識を対象とし、道路標識の視認性に応じてドライバーへ情報提供を行うシステムの実現を目的とする。そのために、車載カメラを用いて道路標識の視認性を推定する技術に注目する。提案手法では、入力された車載カメラ画像から、背景とのコントラスト、道路標識の見え、サイズに基づく複数の画像特徴を統合利用することで道路標識の視認性評価値を計算する。そして、車載カメラ画像系列から計算される一連の視認性評価値を評価することで、時々刻々と変化する走行環境において実際にドライバーが知覚する視認性を高精度に推定することを目指す。実験の結果、1枚の車載カメラ画像のみから道路標識の視認性を推定する従来手法と比較して、車載カメラ画像系列から視認性を推定する提案手法の方が高い視認性推定精度が得られることを確認した。

キーワード 視認性推定, 道路標識, 車載カメラ, 画像特徴, ドライバディストラクション, ITS

1. はじめに

近年、自動車の運転支援システムに対する需要が高まっている。中でも、車載カメラを用いて車両前方の道路標識、歩行者、車両等の物体を自動検知し、その存在をドライバーに知らせる物体検出・提示システムの開発は重要な課題である。このシステムは、車載カメラ画像からの物体検出、検出された物体に関する情報提供という二つの技術からなる。前者の技術に関する研究は従来から数多くなされているものの、それらの多くでは後者の技術までは議論されていない。しかしながら、検出した物体に関する情報をどのようにドライバーに提供するかは重要な問題である。例えば、図1のようなドライバーへの過度の情報提供はドライバーディストラクション（ドライバーの注意散漫）[1]や苛立ちによる不快感の原因になる。そのため、運転支援システムからの情報を適切に取捨選択し、ドライバーが必要とする情報のみを選択的に提供する技術が必要不可欠である。

情報を取捨選択するためのアプローチとして、ドライバーの視線に基づくアプローチが考えられる。例えば、Fletcherらは、ドライバーの視線が向いていない物体の情報を提供するための手法[2]を提案している。しかしながら、視線が物体に向いているということと、その物体を認知しているということとは必ずしも一致しない。従っ

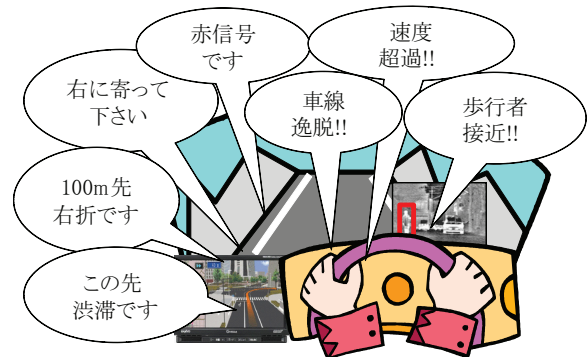


図1 車載システムからドライバーに提供される過量の情報例

て、ドライバーの視線方向のみから運転者に与える情報を選択するだけでは不十分である。一方で、物体の視認性に基づくアプローチも考えられる。例えば、道路標識は交通安全において重要な情報源の一つであるが、その視認性は環境によって大きく変化する。図2(a)では道路標識の視認性が高いため、ドライバーがその存在に気付く可能性が高い。一方、図2(b)では道路標識の視認性が低いため、ドライバーが見落とす可能性が高い。視認性に応じてドライバーへの情報提供の方法を調整できれば、不快感のない運転支援が実現できると考えられる。

これまで我々は、車載カメラ画像から道路標識の視認性を推定する手法を提案してきた[3]。しかし、これまで



(a) 視認性が高いシーン



(b) 視認性が低いシーン

図 2 道路標識の視認性比較

の手法は、1枚の車載カメラ画像を入力とし、ある瞬間における視認性（「瞬時視認性」と呼ぶ）を推定するものであった。一方、実際の交通シーンにおいて、ドライバは瞬時視認性ではなくある一定期間での累積的な視認性（「累積視認性」と呼ぶ）を基に対象の視認性を判断していると考えられる。また、道路標識の視認性は様々な要因によって大きく変化する [4] ため、ある瞬間は視認性が高くても、次の瞬間には低い場合もある。以上のことから本稿では、実際の交通シーンにおいてドライバが感じる視認性を高精度に推定することを目的とし、時系列情報を用いた道路標識の視認性推定手法を提案する。

以降、2. では関連研究について述べ、3. では、提案手法について述べる。続く 4. では、実環境下における提案手法の有効性を評価するための実験について述べ、5. で考察する。最後に 6. でまとめる。

2. 関連研究

物体の見つけやすさには、無意識的な視覚的注意の引かれやすさである誘目性、および、意識的な視覚的注意の引かれやすさである視認性が関係する。以降、誘目性推定および視認性推定に関する研究についてそれぞれ述べる。

2.1 誘目性推定に関する研究

Ittiらは、自然画像を対象として、顕著度画像（saliency map）を用いて入力画像中で誘目性の高い領域を計算す

るモデルを提案している [5]。この顕著度画像は様々な研究で応用され、その有用性が示されている [6, 7]。ただし、無意識的な視覚的注意は被験者の心理状態や知識・興味、事物の予期の有無等に大きく影響を受ける。一方、自動車の運転操作は、外界の状況をすばやく認知し、そこから適切に判断し行動しなければならないという特殊なタスクであるため、ドライバには常に高い負荷がかかる。Ittiらのモデルはそのような状況を想定しておらず、運転中のドライバに対してもそのまま当てはまるとは考えにくい [8]。

2.2 視認性推定に関する研究

車載カメラを用いた道路標識の視認性推定手法は、これまでいくつか提案されてきた [3, 9, 10]。Siegmannらの手法 [9] では輝度情報のみに着目しており、人間の視覚特性が十分に考慮されていない。また、Simonらの手法 [10] では、道路標識の見えをあらかじめ SVM により学習し、SVM における識別関数の出力値を視認性の評価に利用している。しかし、特徴空間における識別境界からの距離と人間からみた視認性は必ずしも一致しない。また、対象物体のみの画像特徴に着目しているため、対象物体とその周辺背景とのコントラストが視認性に及ぼす影響が十分に考慮されていない。そこで我々は、色、エッジ、テクスチャのコントラストに基づく複数の画像特徴を統合利用して道路標識の視認性を推定する手法を提案し、その有効性を確認した [3]。ただし、これまでの我々の手法は1枚の車載カメラ画像から瞬時視認性を推定するためのものであり、視認性の時間的変化を考慮していなかった。実際の交通シーンにおける道路標識の視認性は、照明条件、オクルージョンの程度、道路標識のサイズ等によって時間経過と共に変化する。そのため、ドライバが感じる視認性を高精度に推定するためには、視認性の時間的変化を考慮することが重要である。次節では、車載カメラ画像系列から計算される累積視認性を基に道路標識の視認性を推定する手法を提案する。

3. 提案手法

図3に提案手法の処理の流れを示す。提案手法では、各時刻 t において入力される車載カメラ画像から、道路標識の瞬時視認性を推定する。そして、車載カメラ画像系列に対して推定された一連の瞬時視認性を統合し評価することで累積視認性を推定する。なお、提案手法では、各入力画像上での道路標識の位置、サイズ、種類は既知であるとする。車載カメラ画像から道路標識を高精度に検出・認識する手法はこれまでも数多く提案されており [11, 12]、本研究でもこれらが適用可能であると考えられる。以降、各ステップについて詳述する。

3.1 瞬時視認性の推定

一般に対象の視認性には、前景と背景の色、エッジ、

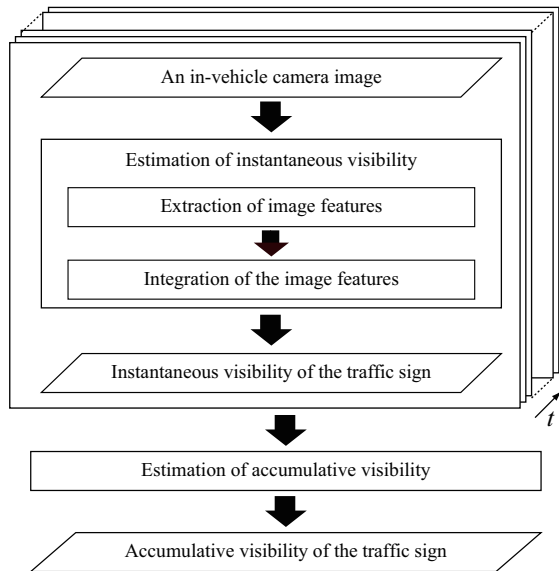


図3 提案手法における処理の流れ

テクスチャのコントラスト、状態（品質、オクルージョンの程度）、サイズ等が関係する。本手法では、それら を評価する5つの画像特徴量 f_i ($i = 1, \dots, 5$) を統合利用することにより、道路標識の瞬時視認性を推定する。以下に、各特徴量の計算方法を述べる。

3.1.1 画像特徴量の計算

道路標識とその周辺のコントラストに基づく特徴量 f_i ($i = 1, 2, 3$) を以下のようにして計算する [3]。まず、道路標識とその周辺を含んだ画像（「標識周辺画像」と呼ぶ）を入力画像 I から切り出す。次に、図4に示すように、標識周辺画像を標識領域 s と複数の部分背景領域 $b_n \in \mathbb{B}$ に領域分割する。これを元に、それぞれ以下のように定義される色、エッジ、テクスチャのコントラスト $c_i^{(b_n)}$ ($i = 1, 2, 3$) を計算する。

- 色コントラスト $c_1^{(b_n)}$: 標識領域 s と部分背景領域 b_n における平均色のユークリッド距離
 - エッジコントラスト $c_2^{(b_n)}$: 標識領域 s と部分背景領域 b_n における平均エッジ強度の差
 - テクスチャコントラスト $c_3^{(b_n)}$: 標識領域 s と部分背景領域 b_n それぞれにおける正規化色ヒストグラム間のヒストグラムインターセクションに基づく距離
- その後、各 $c_i^{(b_n)}$ を用いて特徴量 f_i ($i = 1, 2, 3$) を次式で計算する。

$$f_i = \sum_{b_n \in \mathbb{B}} \frac{a^{(b_n)}}{a^{(\mathbb{B})}} c_i^{(b_n)} \quad (1)$$

ここで、 s の重心を原点とする画素の座標を p として、 $a^{(b_n)} = \sum_{p \in b_n} \|p\|^{-1}$ 、 $a^{(\mathbb{B})} = \sum_{b_n \in \mathbb{B}} a^{(b_n)}$ である。

道路標識の状態に基づく特徴量 f_4 を次式で計算する。

$$f_4 = S(s, s_t) \quad (2)$$

ここで、 s_t は見えの劣化を一切含まない理想的な標識テンプレートであり、 S は標識領域 s と標識テンプレート

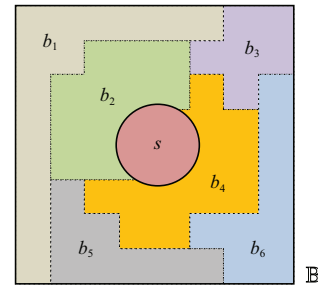


図4 標識周辺画像における標識領域 s と部分背景領域 $b_n \in \mathbb{B}$ ($n = 1, \dots, 6$) の例

s_t 間の SSD (Sum of Squared Difference) に基づく類似度である。

道路標識のサイズに基づく特徴量 f_5 を次式で計算する。

$$f_5 = \frac{A^{(s)}}{A^{(I)}} \quad (3)$$

ここで、 $A^{(s)}$ 、 $A^{(I)}$ はそれぞれ標識領域 s 、入力画像 I の面積である。

3.1.2 画像特徴の統合

特徴量 $f = (f_1, \dots, f_5)$ を元に、ある時刻における瞬時的な視認性評価値 \hat{v} を次式で計算する。

$$\hat{v} = \mathbf{w}^T \phi(\mathbf{f}) = \sum_{z=1}^Z w_z \phi_z(\mathbf{f}) \quad (4)$$

ここで、 $\mathbf{w} = (w_1, \dots, w_Z)^T$ は基底関数を要素にもつベクトル $\phi(\mathbf{f}) = (\phi_1(\mathbf{f}), \dots, \phi_Z(\mathbf{f}))^T$ の重みである。

3.2 累積視認性の推定

各時刻 t で計算される $\hat{v}^{(t)}$ を次式で統合することで、累積視認性の評価値 \hat{V} を計算する。

$$\begin{aligned} \hat{V} &= \frac{1}{T_p} \sum_{t=0}^{T_p-1} \hat{v}^{(\tau-t)} \\ &= \frac{1}{T_p} \sum_{t=0}^{T_p-1} \sum_{z=1}^Z w_z \phi_z(\mathbf{f}^{(\tau-t)}) \\ &= \sum_{z=1}^Z w_z \left[\frac{1}{T_p} \sum_{t=0}^{T_p-1} \phi_z(\mathbf{f}^{(\tau-t)}) \right] \\ &= \mathbf{w}^T \Phi \end{aligned} \quad (5)$$

ここで、 τ は現在時刻、 T_p は入力画像の総数であり、 Φ は次式で定義される。

$$\Phi = \frac{1}{T_p} \left(\sum_{t=0}^{T_p-1} \phi_1(\mathbf{f}^{(\tau-t)}), \dots, \sum_{t=0}^{T_p-1} \phi_Z(\mathbf{f}^{(\tau-t)}) \right)^T \quad (6)$$

提案手法では、 \hat{V} の値が高いほど視認性が高く、 \hat{V} の値が低いほど視認性が低いと判定する。

表 1 本実験で対象とする道路標識の例

種類	含まれる道路標識
警戒標識	
規制標識	
指示標識	

4. 実験

被験者実験を通して提案手法の有効性を評価した。本実験では、車載カメラ画像 1 枚を用いて瞬時視認性を評価する従来手法 [3] を比較手法として採用した。また、道路交通において特に重要と考えられる表 1 に示す道路標識を視認性推定の対象とした。以降、実験準備、評価条件、実験結果を順に述べる。

4.1 実験準備

評価用セットおよびパラメータ w の学習用セットを以下のような手順で作成した。まず、日中の様々な天候の下で市街の道路標識を車載カメラ (1920 × 1080 pixels, 15 fps) で撮影した。次に、撮影された車載カメラ映像の中から、フレーム内に道路標識全体が含まれる映像区間を $N = 100$ 本 (19 ~ 169 frames) 抽出し、これを評価用セットとした。同様に、評価用セットとは異なる映像区間から、それぞれ異なるシーンにおける $M = 59$ フレームを抽出し、これを学習用セットとした。

評価用映像および学習用画像に対する視認性評価値の ground-truth は、以下に述べるような被験者実験 (被験者: 20 ~ 30 代の男女 8 名) により決定した。まず、コンピュータのディスプレイを用いて、被験者に評価用映像区間を 1 回だけ提示した。次に、被験者に対象標識の視認性を [0,1] で絶対評価させた。このとき、対象とする道路標識が見つけれなかった場合は評価値を 0 とした。そして、全被験者による平均評価値をその映像区間に対する ground-truth として利用した。上記手続きを全ての評価用映像区間に対して行い、評価用セットに対する視認性評価値の ground-truth V_n ($n = 1, \dots, N$) を得た。同様に、学習用セットに対する視認性評価値の ground-truth U_m ($m = 1, \dots, M$) を得た。

4.2 評価条件

提案手法および従来手法では、コントラスト $c_i^{(b_n)}$ の計算には RGB 色空間を利用し、基底関数 $\phi(f)$ として交差項を含む 2 次の多項式基底関数 ($Z = 20$) を利用した。これらのパラメータは実験的に決定した。また、

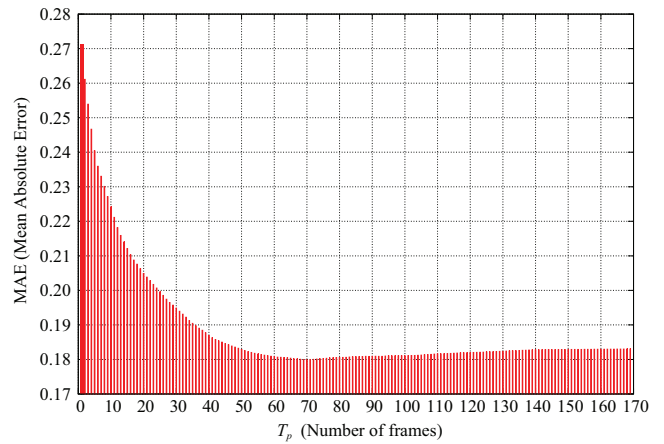


図 5 実験結果: 従来手法 ($T_p = 1$) および提案手法 ($T_p > 1$) による MAE

学習用セットに対する視認性評価値の ground-truth U_m ($m = 1, \dots, M$) を用いた回帰により w を求めた。

視認性推定精度の評価基準としては、次式で示す MAE (Mean Absolute Error) を使用した。

$$\text{MAE} = \frac{1}{N} \sum_{n=1}^N |V_n - \hat{V}_n| \quad (7)$$

ここで、MAE は [0,1] の範囲の値をとり、低いほど視認性推定の精度が高いことを意味する。提案手法では、累積視認性の評価値の計算に使用するフレーム数 T_p を 1 ~ 169 (評価用セットにおける映像区間の最長フレーム数) の間で 1 ずつ変化させながら MAE を評価した。このとき、 T_p より短い映像区間に対しては、その映像区間内での平均評価値を累積視認性の評価値とした。従来手法では、評価用セットにおける各映像区間の最終フレーム (道路標識がフレームアウトする直前のフレーム) に対する視認性評価値から MAE を評価した。なお、提案手法における $T_p = 1$ の場合、従来手法と等価である。

4.3 実験結果

従来手法 ($T_p = 1$) および提案手法 ($T_p > 1$) による MAE を図 5 に示す。従来手法による MAE は 0.271 であった。一方、提案手法では、 T_p によらず従来手法よりも MAE が低く、特に $T_p = 70$ のときに MAE は 0.180 と最も高精度に視認性を推定できた。なお、従来手法と提案手法 ($T_p = 70$) の間の MAE の差に関して t 検定を行った結果、有意水準 1% で有意差が確認された。これらのことから、車載カメラ画像系列を用いて道路標識の視認性を推定する提案手法の有効性を確認した。

5. 考察

提案手法の有効性、および、シーンコンテキストの考慮の必要性に関して考察を述べる。

5.1 提案手法の有効性

まず、車載カメラ画像系列を用いて道路標識の視認性

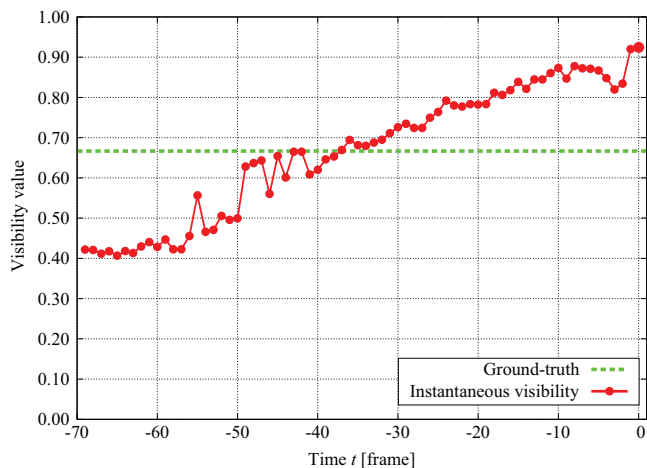


図 6 瞬時視認性の推移の例

を推定することの有効性に関して、ある評価用映像区間に対する最終 70 フレームにおける瞬時視認性の評価値の推移を図 6 に示す。また、その映像区間に含まれるフレームの例を図 7 に示す。最終フレーム ($t = 0$) のみから視認性を推定する従来手法の視認性評価値は 0.92 であった。一方、最終フレームから 70 フレームを用いて視認性を推定する提案手法の視認性評価値は 0.67 であり、ground-truth とほぼ一致した。なお、この映像区間においては、 $t = -42$ 周辺において瞬時視認性の評価値と ground-truth がほぼ一致していた。しかし、他の映像区間における瞬時視認性の評価値の推移を確認すると、瞬時視認性の評価値が ground-truth と一致するタイミングには規則性は確認できなかった。従って、道路標識の視認性を瞬時視認性から推定することは困難であると考えられる。これに対して、提案手法は車載カメラ画像系列を用いて平均的な瞬時視認性を評価するため、時々刻々と変化する走行環境においても道路標識の視認性を比較的安定して推定できると考えられる。このことが、提案手法により最も小さい MAE (式 (7)) が得られた理由の一つであると考えられる。

また、提案手法で利用する画像特徴の有効性に関して、道路標識の状態に基づく画像特徴 f_4 およびサイズに基づく画像特徴 f_5 を利用せずに視認性を推定した場合の MAE は 0.19 であった。このことから、道路標識の状態やサイズに基づく画像特徴の利用が道路標識の視認性推定に有効であることが確認された。

5.2 シーンコンテキスト考慮の必要性

実際の走行シーンにおいては、他の道路標識、看板、信号機、歩行者、車両の存在等によって、道路標識の見つけやすさは大きく影響されると考えられる。実際に、被験者らによる視認性評価値のばらつきは、周囲に道路標識以外の物体が存在しないシーンでは小さく、道路標識以外の様々な物体が周囲に存在するシーンでは大きいという傾向がみられた。例えば、図 8(a) のようなシーン



(a) 時刻 $t = -70$ [フレーム]



(b) 時刻 $t = -42$ [フレーム]



(c) 時刻 $t = 0$ [フレーム] (最終フレーム)

図 7 図 6 に対応する評価用映像区間におけるフレームの例

では、ドライバの注意を引きやすい目立つ物体が存在しない。このような単純なシーンコンテキストにおいては、道路標識の視認性は局所的な画像特徴のみで評価できると考えられる。しかし、図 8(b) のようなシーンでは、複数の道路標識、自転車とそれを追い越そうとする前方車両等が存在する。このように複雑なシーンコンテキストにおいては、局所的な画像特徴のみで対象標識の視認性を評価することは困難である。よって、更に高精度な視認性推定を実現するためには、局所的な画像特徴だけでなくシーンコンテキストを考慮した大局的な画像特徴の考慮が必要であると考えられる。

6. まとめ

本論文では、不快感のない運転支援システムの実現に向けた取り組みの一つとして、時間的環境変化を考慮し



(a) 単純なコンテキストのシーン



(b) 複雑なコンテキストのシーン

図 8 コンテキストの複雑さが異なるシーンの比較

た道路標識の視認性推定手法を提案した。提案手法では、実際の走行シーンにおいてドライバーが感じる視認性を高精度に推定するため、車載カメラ画像系列を用いて計算される瞬時視認性の評価値を統合利用する。被験者実験の結果、1枚の車載カメラ画像から視認性を推定する従来手法よりも車載カメラ画像系列を用いて視認性を推定する提案手法の方が高精度に視認性を推定できることを確認した。今後は、更に高精度な視認性推定を実現するため、シーンコンテキストを考慮した大局的な画像特徴を検討していく。

謝辞 本研究の一部は、戦略的創造研究推進事業 CREST，科学研究費補助金，日本学術振興会特別研究員奨励費による。また，本研究では画像処理に MIST ライブラリ^(注1)を使用した。

文 献

- [1] M. Pettitt, G. Burnett, and A. Stevens, "Defining driver distraction," Proc. 12th World Cong. on Intelligent Transportation Systems, pp.1-12, November 2005.
- [2] L. Fletcher, G. Loy, N. Barnes, and A. Zelinsky, "Correlating driver gaze with the road scene for driver assistance systems," Robotics and Autonomous System, vol.52, no.1, pp.71-84, July 2005.
- [3] K. Doman, D. Deguchi, T. Takahashi, Y. Mekada, I. Ide, H. Murase, and Y. Tamatsu, "Estimation of traffic sign visibility toward smart driver assistance," Proc. 2010 IEEE Intelligent Vehicles Symposium, pp.45-50, June 2010.

- [4] Commission Internationale de l'Eclairage (CIE), "The conspicuity of traffic signs in complex backgrounds," CIE Tech. Rep. 137, 2000.
- [5] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," IEEE Trans. Pattern Analysis and Machine Intelligence, vol.20, no.11, pp.1254-1259, November 1998.
- [6] K. Miyazato, A. Kimura, S. Takagi, and J. Yamato, "Real-time estimation of human visual attention with dynamic Bayesian network and MCMC-based particle filter," Proc. 2009 IEEE Int. Conf. on Multimedia and Expo, pp.250-257, June 2009.
- [7] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," Proc. 2009 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, pp.1597-1604, June 2009.
- [8] L. Simon, J.P. Tarel, and R. Brémond, "A new paradigm for the computation of conspicuity of traffic signs in road images," Proc. Int. Conf. of the 26th Session of the CIE, vol.2, pp.161-164, July 2007.
- [9] P. Siegmann, S. Lafuente-Arroyo, S. Maldonado-Bascón, P. Gil-Jiménez, and H. Gómez-Moreno, "Automatic evaluation of traffic sign visibility using SVM recognition methods," Proc. 5th WSEAS Int. Conf. on Signal Processing, Computational Geometry & Artificial Vision, pp.170-175, September 2005.
- [10] L. Simon, J.P. Tarel, and R. Brémond, "Alerting the drivers about road signs with poor visual saliency," Proc. 2009 IEEE Intelligent Vehicles Symposium, pp.48-53, June 2009.
- [11] K. Doman, D. Deguchi, T. Takahashi, Y. Mekada, I. Ide, and H. Murase, "Construction of cascaded traffic sign detector using generative learning," Proc. 4th Int. Conf. on Innovative Computing, Information and Control, ICICIC-2009-1362, December 2009.
- [12] H. Ishida, T. Takahashi, I. Ide, Y. Mekada, and H. Murase, "Generation of training data by degradation models for traffic sign symbol recognition," IE-ICE Trans. on Information and Systems, vol.E90-D, no.8, pp.1134-1141, August 2007.

(注1): <http://mist.murase.m.is.nagoya-u.ac.jp/>