

局所的特徴と大局的特徴の統合に基づく歩行者の視認性定量化

若山 雅史[†] 道満 恵介[†] 出口 大輔[†] 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

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

E-mail: [†]{mwakayama,kdoman,ddeguchi,ide,murase}@murase.m.is.nagoya-u.ac.jp

あらまし 本稿では、運転時に見落とししやすい情報のみをドライバに提示するシステムに着目し、歩行者の視認性を定量化する手法を提案する。近年、事故防止を目的とした多種多様な安全運転支援システムが開発されており、ドライバへ様々な情報提示が可能になってきている。しかしながら、ドライバへの過度の情報提示はドライバの注意力低下を招くおそれがあり、情報の適切な取捨選択技術が必要とされている。この問題の解決法の1つとして、ドライバが認識し難いと予想される情報のみを提示する方法が考えられる。物体の認識しやすさはその物体の見え、すなわち視認性と密接に関係している。そこで本研究では、ドライバの見落としによる事故が多いとされる歩行者を対象とし、検出した歩行者情報の取捨選択の基準として視認性を利用することを考える。本稿では、視覚探索のモデルを参考に性質の異なる2つの特徴量を統合することで視認性を定量化する手法を提案する。対比較に基づく被験者実験により、提案手法の視認性評価値は最大で97%の一致率を示した。

キーワード 視認性定量化, 歩行者, 車載カメラ, 画像特徴, 顕著性マップ, ITS

Quantification of the Visibility of Pedestrians Based on Integration of Local and Global Features

Masashi WAKAYAMA[†], Keisuke DOMAN[†], Daisuke DEGUCHI[†],

Ichiro IDE[†], Hiroshi MURASE[†], and Yukimasa TAMATSU^{††}

[†] Graduate School of Information Science, Nagoya University, Nagoya, Aichi, Japan

^{††} DENSO CORPORATION, Kariya, Aichi, Japan

E-mail: [†]{mwakayama,kdoman,ddeguchi,ide,murase}@murase.m.is.nagoya-u.ac.jp

Abstract This paper proposes a method to quantify the visibility of pedestrians using in-vehicle camera. Since many driver assistance systems have been developed, various information can be presented to the drivers. If too much information is presented, their attention may decrease due to the limitation of processing ability of the driver. Therefore, mechanisms for providing appropriate information are needed. One of the solution to this problem is to present drivers only information that is difficult to recognize. And it is considered that the visibility is one of the most important factor for recognition. Therefore, we focus on estimating the visibility of pedestrians to avoid pedestrian-car accidents. In this paper, we calculate two different features based on a visual search model, then combine them to evaluate the visibility of pedestrians. And we evaluate the performance of the proposed method. The experimental result showed that the proposed method could obtain 97 percent at maximum in coincidence rate.

Key words Quantification of the visibility, Pedestrian, In-vehicle camera, Image feature, Saliency map, ITS

1. はじめに

近年、ABSの搭載や車体の設計技術の向上などにより、自動車の安全性能は飛躍的に向上した。また、運転中に安全面での

支援を行うシステムも実用化されてきている。例えば、歩行者の検出システム [1]、車線維持 [2] や車間維持 [3] などがある。こうした自動車に関する安全技術の進歩により、交通事故の発生件数は近年減少傾向にある。しかし、依然として交通事故に

よる死亡者数は年間約5千人に上る。そのうちの約3割が歩行者であり、事故原因の多くは前方不注意や安全確認を怠ったことによるドライバの認知遅れ、認知ミスである。よって、歩行者の存在を自動で検出し、それに関する情報を提示するシステムは、安全運転支援に有効である。しかし、これまでの歩行者の検出技術に関する研究では、どのようにドライバへ情報提示を行うかが検討されていない [4, 5]。運転操作は常に適切な外界の認知と判断を必要とするため、過度の情報提示はドライバの苛立ちや注意力低下を招くおそれがある [6]。そのため、検出した情報をドライバや周囲の状況に応じて取捨選択する技術が必要であり、ドライバが認識し難い情報のみ提示されることが好ましい。

Fletcher らは視線計測システムを利用し、視線の動きからドライバが知覚していないと思われる情報を提示する手法を提案している [7]。しかし、ドライバの視線が向いている対象と、実際に知覚している対象は必ずしも一致しない。つまり、“注視していても知覚していない状態 [8]”と、“知覚していても注視しない状態”が起こり得るため、ドライバの視線情報のみによる提示の有無の判断では不十分である。また、ドライバの脇見や心理状態を推定 [9, 10] し、情報の取捨選択基準とする方法も考えられる。しかし、通常の心理状態でも歩行者の知覚しやすさはその周囲環境により異なるため、この手法単体では不十分である。

一方、“ドライバから見た対象の知覚しやすさ”を測る指標として視認性を利用した手法が考えられる。例えば歩行者の視認性に関して、車載カメラから撮影した異なるシーンの歩行者を図1に示す。それぞれの画像には歩行者が1人ずつ撮影されているが、明るさや周囲の状況によってその視認性は大きく異なる。ドライバが見落としやすい、つまり視認性が低いと考えられる歩行者のみを警告すれば、見落としによる事故の危険を低下させると共に、ドライバの情報過多の防止が実現できる。よって本研究では、ドライバからの見え方に近い画像が得られる車載カメラで歩行者を撮影した画像を使用し、視覚探索のモデルに基づいた歩行者の視認性の定量化することを考える。

本稿では、人の視覚探索のモデルに基づいて計算される、性質の異なる2つの特徴量を統合することで歩行者の視認性を定量化し、被験者実験によりそれが人の感覚にどの程度一致するか評価する。以降、2節では視覚探索のモデルおよび関連研究について述べ、3節では、歩行者の視認性を定量化する手法を提案する。続く4節では、視認性評価値の妥当性を評価するための被験者実験について述べる。最後に、5節でまとめと今後の課題について述べる。

2. 視覚探索のモデル

視認性について考える上で、まず視覚探索のプロセスを考慮することは重要である。認知心理学において、あるシーンから対象物体（ターゲット、target）を探索するタスクは visual search と呼ばれ、それに関連する研究が文献 [11] にまとめられている。古典的には、visual search は全ての（あるいは多くの）物体を並列的に処理するプロセス（parallel search）と、



(a) 歩行者の視認性：高



(b) 歩行者の視認性：低

図1 異なるシーンにおける歩行者の視認性の違い

一度に1つの物体のみに注意を向けるプロセス（serial search）の相互作用によって成立すると考えられている。このような過去の研究から得られた知見に基づき、Wolfe は Guided Search と呼ばれる visual search のモデルを提案している [11]。このモデルは以下の2つのステージから構成される。

1. 複数の物体から、“ガイダンス”と呼ばれるものの作用により、探索に重要そうな物体、あるいは領域が並列的に選択される（parallel search の性質）。
2. それらに対し“認識処理”が行われ、対象物であるか非対象物（ディストラクタ、distractors）であるかが分かった時点で逐次的に注意が向けられる（serial search の性質）。

以降の説明のため、重要そうな物体／領域のことをターゲット候補と呼ぶものとする。ガイダンスは「ターゲットは赤い」などの知識に基づく top-down ガイダンスと、「視覚的に目立つ」

などの知識に基づかない bottom-up ガイダンスに分けられ、全体から素早くターゲット候補を絞るために用いられる。また、認識処理とは各ターゲット候補の情報を対象／非対象のしきい値を超えるまで蓄積する処理であり、ある程度並列的に行われる。加えて、一般にディストラクタの数（セットサイズ, set size）が多い画像では、そうでない画像に比べて認識処理により多くの時間を要するため [12], ターゲットは見つけにくくなる。従って、視認性定量化には

- (1) parallel search (ガイダンスの効果)
- (2) serial search
- (3) ディストラクタの数による探索難易度の変化

の3点が考慮されている必要があると考えられる。以下にそれぞれの関連研究について述べる。

(1) **parallel search (ガイダンスの効果)**: bottom-up ガイダンスを考慮する手法として, Itti らの顕著性マップ (Saliency map) [13] がある。顕著性マップとは, 画像の各部分への視覚的注意の向きやすさ (顕著度) を表現するマップである。Itti らの手法は, 入力画像全体から輝度・色・エッジなどの基本的な画像特徴を抽出し, それらの空間的コントラストを顕著度として計算する。一方, Wolfe は画像特徴の空間的コントラストを bottom-up ガイダンスとして利用している。よって, Itti らの顕著性マップは bottom-up ガイダンスの性質を持ち, 顕著度が高い領域はガイダンスの影響によりターゲット候補として選択される可能性が高いと考えられる。なお, Navalpakkam らは top-down ガイダンスの性質を取り入れた顕著性マップについて検討しているが [14], その性質は十分に解明されていない。

(2) **serial search**: Paulmier らは, 歩行者のシルエットとその周辺背景の輝度をもとに視認性評価値を定義し, 被験者実験により歩行者の見えと視認性の関係を明らかにしている [15]。実験室環境における被験者実験の結果から, 輝度だけでなく背景の複雑度も視認性に影響することを示している。また木村らは, 人間の視覚的注意に影響を及ぼす要因を整理し, 交通信号機の視認性定量化に有効な画像特徴の検討を行い, エッジ強度に基づく複雑度が視認性定量化に有効であることを示した [16]。これらの研究は, 対象とその周辺背景の画像特徴の関係に着目している。つまりターゲット周辺のごく一部のみを考慮しているため, serial search の性質を持つと考えられる。

(3) **ディストラクタの数による探索難易度の変化**: 今までこの性質を視認性の定量化に利用した研究は行われていない。

以上の (1) ~ (3) をふまえ, 本研究ではこれらの手法を組み合わせた歩行者の視認性定量化手法を提案する。以降の説明のため, 対象と背景の画像特徴の関係を“局所的特徴”と呼び, ガイダンスを考慮したターゲット候補間の関係を“大局的特徴”と呼ぶものとする。前者は (2) に対応し, 後者は (1) および (3) に対応するものである。次節では具体的な視認性定量化手法について述べる。

3. 提案手法：歩行者の視認性定量化手法

本稿では車載カメラ画像を用いた歩行者の視認性推定手法を

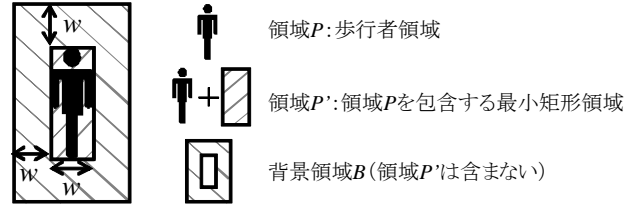


図2 局所特徴量の計算で使用する歩行者と背景領域の定義

提案する。本手法では, まず局所の特徴と大局の特徴を定義し, それぞれの特徴量を独立に定量化する。局所の特徴量として, 歩行者と背景の画像特徴の差を利用する。また大局の特徴量として, 画像中のターゲット候補領域と真のターゲット (歩行者) 領域の顕著度の差の総和を利用する。最後に, これら2つの値を統合して, 最終的な歩行者の視認性評価値を計算する。以降, 各特徴量の計算, および統合の手法について順に述べる。

3.1 局所の特徴量の計算

単位面積あたりに含まれるエッジの量は複雑度を表現する特徴量の一つであり, 画像の複雑度は対象物の視認性に関する [15, 16]。複雑度の差が大きいほどターゲットは見つけやすいと考えられる。

そこで, 図2に示す領域に対し Sobel フィルタを適用し, 歩行者領域 P の平均エッジ強度 E_p と背景領域 B の平均エッジ強度 E_b を計算する。なお, 歩行者領域は検出と同時に輪郭を抽出する手法 [17] などにより抽出できる (本稿では手動で行う)。その後, 次式により複雑度の差に基づく局所の特徴量 V_l を計算する。

$$V_l = |E_p - E_b| \quad (1)$$

3.2 大局の特徴量の計算

歩行者領域 P と, P を除くターゲット候補領域集合 \mathbb{D} の顕著度の差を用いて大局の特徴量を計算する。ディストラクタの数 (ここでは \mathbb{D} の数に等しい) を考慮するため, 総和を用いて次式により計算する。

$$V_g = - \sum_{D_i \in \mathbb{D}} (A(D_i) - A(P)) \quad (2)$$

ただし,

$$A(X) = w(X) \sum_{x \in X} S(x) \quad (3)$$

であり, $S(x)$ は領域 X 中の各画素 x の顕著度の値である。また, $w(X)$ は領域の面積による見えの変化を考慮する係数であり, 次式で表される。

$$w(X) = \frac{\log |X|}{|X|} \quad (4)$$

$|X|$ は領域 X の画素数である。

次に \mathbb{D} の抽出について述べる。顕著度が高い領域ほどガイダンスの影響によりターゲット候補として選択される可能性が高いと考えられるため, 以下で述べるように顕著性マップを用いて近似的に抽出する。一連の流れを図3に示す。

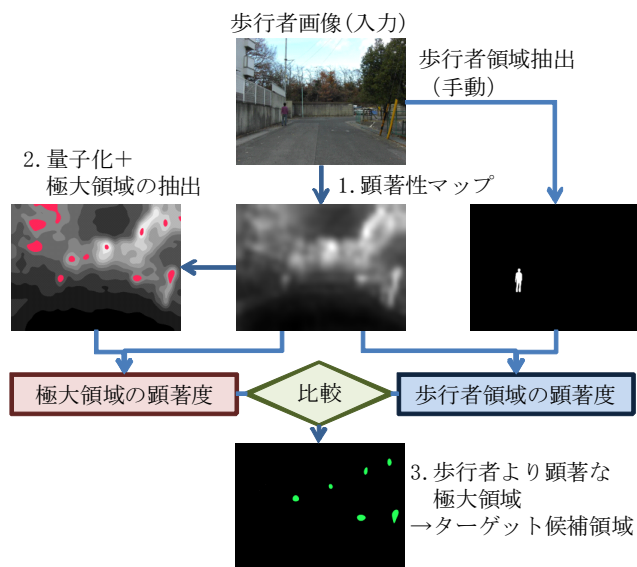


図3 顕著性マップを利用した、歩行者を除くターゲット候補領域の抽出

1. 入力画像から顕著性マップを作成し、入力画像と同じ解像度に拡大する。
2. 顕著性マップを均等量子化し、量子化した顕著度が極大値をとる隣接領域のみを残す。ただし、隣接領域の画素数がある一定数（ここでは200画素）以下であった場合は、周囲のより小さな（量子化後の）顕著度を持つ隣接領域に統合する。
3. 残った各領域中の平均顕著度が歩行者領域の平均顕著度より高い領域をDとする。これは歩行者より平均顕著度が低い領域は歩行者よりも先に認識されないことを仮定している。

3.3 特徴量の統合

3.1 および3.2で計算された特徴量を次式で統合する。

$$\text{Visibility} = (1 - \alpha)V_l + \alpha V_g \quad (5)$$

α は統合する特徴量の比を表すパラメータ ($0 \leq \alpha \leq 1$)である。本手法では Visibility の高低により、視認性の高低を評価する。

4. 評価実験

前節で述べた視認性評価値の有効性を確認するために被験者実験を行った。以降でその手順と結果について述べる。

4.1 実験準備

背景の複雑さ、歩行者の見えなどにより視認性は変化するため、背景5種、服の色4種、カメラからの距離3種をそれぞれ変化させた合計60枚の車載カメラ画像を用意した。これらを次に説明する被験者実験で利用した。なお、撮影した画像全てにおいて、写っている歩行者は1人だけであり、全て同一人物である。

4.2 実験手順

まず、収集した60枚の画像を、主観的に歩行者が見つけれ

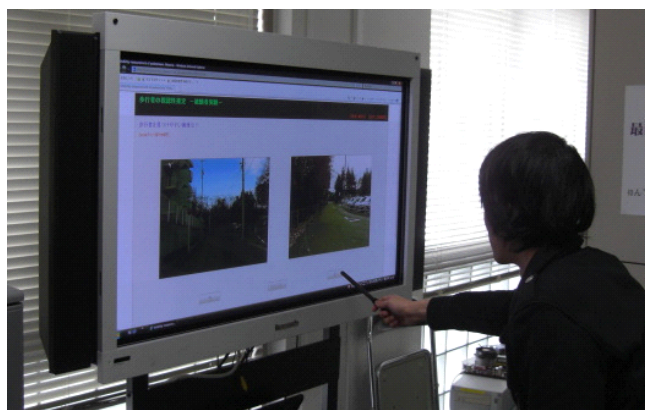


図4 被験者実験の様子

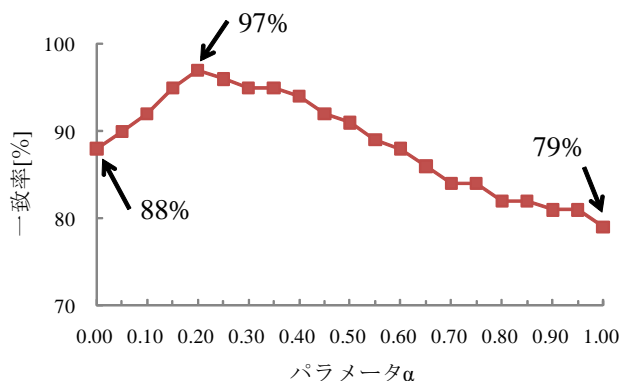


図5 パラメータ α に対する一致率の変化

すいものを見つけにくいものに分類する。それらの中から10枚ずつ、前述の視認性の変化要因に偏りがないように選択し、両者を並べて $10 \times 10 = 100$ 対の画像対を用意する。これをランダムに並べ替えたものを被験者に提示し、「歩行者が見つかりやすいのはどちらか」を「左」、「右」、「わからない」のいずれかで回答を得る。10名の被験者に対して実験を行い、合計1000回の回答を得た。なお上記手続きは、すべての被験者で図4のような同一のディスプレイ、およびインターフェースを利用した。

評価に際しては、式(5)で計算される視認性評価値 Visibility の大小関係と、被験者からの回答の多数決による大小関係が一致する画像対を一致画像対、そうでないものは不一致画像対と定義し、次式で計算される一致率を利用した。

$$\text{一致率} = \frac{\text{一致画像対の数}}{\text{全画像対の数}} \times 100[\%] \quad (6)$$

なお、「わからない」という回答は被験者からの回答として計数しなかった。

4.3 実験結果

統合のパラメータ α を0から1の範囲で変化させ、一致率を評価した。一致率の変化のグラフを図5に示す。図から、一致率は $0 < \alpha < 1$ の範囲で1つの極大値をとるように変化したことがわかる。式(5)から、提案手法の視認性評価値は α を0に設定すると局所の特徴量に等しく、 α を1に設定すると大局的の特徴量に等しい。 $\alpha = 0.2$ のとき、一致率が最も高い97%を示した。これは、局所の特徴量のみを考慮 (α を0に設定) した88%より9%高い。使用した画像対の数は100であるため、局

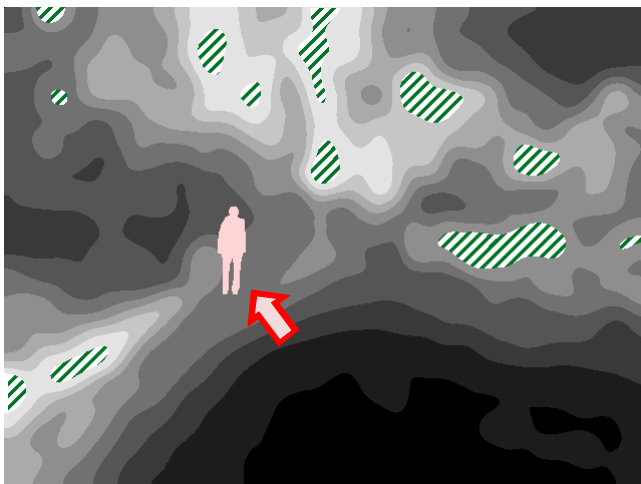


(a) 被験者による視認性評価：高

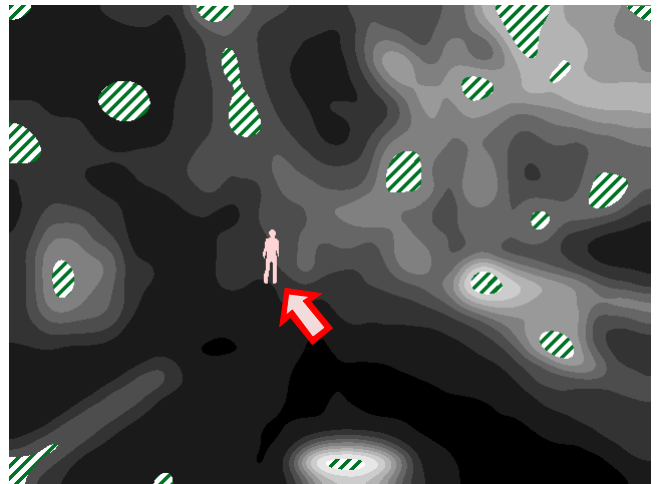


(b) 被験者による視認性評価：低

図 6 特徴量の統合により不一致画像対から一致画像対となった例：局所的特徴量のみを考慮 ($\alpha = 0$ に設定) した際の評価値は (a) < (b) であったが、大局的特徴量と統合 ($\alpha = 0.2$ に設定) した際の評価値は (a) > (b) となった。



(a) 図 6(a) から抽出されたターゲット候補領域



(b) 図 6(b) から抽出されたターゲット候補領域

図 7 図 6 の各歩行者画像から顕著性マップを用いて抽出されたターゲット候補領域 (斜線部) : (b) は歩行者領域 (矢印の先) の顕著度が低いため、比較的顕著度が低い極大領域もターゲット候補領域として抽出されている。結果として (b) の大局的特徴量は小さくなる。

所的特徴量のみを考慮した際の不一致画像対のうち、少なくとも 9 画像対で視認性評価値の大小関係が逆転したことになる。実際には、11 画像対で視認性評価値の大小関係が逆転し、そのうちの 10 画像対が不一致から一致となった (残りの 1 画像対は一致から不一致となった)。

5. 考 察

本実験において被験者の回答との一致率は $\alpha = 0.2$ のとき最高の 97% となった。これは、局所的特徴量または大局的特徴量のどちらか一方のみで視認性を評価するよりも、両者を統合して視認性を評価した方が一致率が上昇することを示している。

しかし、本実験において使用した車載カメラ画像の数は 20 枚であり、実際の背景や服装等のバリエーションを考えると十分とは言えない。そのため、バリエーションを増やしても同様の効果が得られるか検討する必要があると考えられる。本実験は統合に基づく視認性の定量化が効果的に働く可能性を示したといえる。

従来手法 [15,16] と同様の考えに基づく局所的特徴量による視認性定量化手法と、統合に基づく視認性定量化手法を比較する。 $\alpha = 0$ では不一致画像対であったが、 $\alpha = 0.2$ では視認性評価値の大小関係が逆転し、一致画像対となった例を図 6 に

示す。また、図 6 のそれぞれの歩行者画像から抽出されたターゲット候補領域を図 7 に示す。局所的特徴量、すなわち歩行者と背景の複雑度の差は図 6(a) よりも図 6(b) の方が大きい値を示した。しかし、図 7(b) から、対応する図 6(b) の歩行者の顕著度は他の領域に比べて小さく、顕著度があまり高くないと思われる領域がターゲット候補領域として抽出されている。この原因としては、図 6(b) の歩行者とその周辺の背景の色・輝度などの空間的コントラストが小さいことが考えられる。結果として図 6(b) の大局的特徴量は図 6(a) よりも小さくなり、統合が評価値の大小を逆転するように働いたと考えられる。このように、統合に基づく手法が有効となるためには、局所的特徴量と大局的特徴量が互いの欠点を補う必要があると考えられる。このような性質を持つ特徴量については今後さらに検討する余地がある。

また、本手法では top-down ガイダンスを考慮しなかった。対比較による実験はその性質上画像対をじっくりと観察することができる。そうした状況では局所的特徴量のみで視認性の高低が判断されてしまい、ガイダンスを考慮する効果が薄れると考えられる。今後は、例えば歩行者がいると判断するまでの反応時間を計る、といった方法で評価するとともに、大局的特徴量に top-down ガイダンスを取り入れる必要があると考えられる。

6. ま と め

本稿では、事故防止に期待されている安全運転支援システムとして歩行者検出システムを取り上げ、情報過多を防止する方法として視認性の高低に着目した。提案手法では、視覚探索のモデルに基づき、歩行者と背景の画像特徴の関係に基づく局所的特徴量と、ガイダンスの考慮に基づく大局的特徴量の統合により視認性を定量化した。被験者実験では、背景や歩行者の色を変化させて作成した車載カメラ画像を利用し、対比較実験から得られた一致率により視認性を評価した。その結果、提案手法により最大で 97% の一致率が得られることを確認した。局所的特徴量のみを用いた場合は 88%、大局的特徴量のみを用いた場合は 79% であった。これらから、本稿で提案した局所的特徴量と大局的特徴量の統合に基づく視認性評価値が人間の感覚に最もよく一致することを確認した。今後の課題としては、車載カメラ映像を対象とした歩行者の視認性定量化手法の検討、top-down ガイダンスを考慮した大局的特徴の導入、より多くの車載カメラ画像での評価、実環境に近い状況下での実験等が挙げられる。

謝辞 日頃より研究のご指導を頂いている名古屋大学村瀬研究室諸氏に深く感謝する。本研究の一部は、JST 戦略的創造研究推進事業 CREST および科学研究費補助金による。また、本研究では画像処理に MIST ライブラリ (<http://mist.murase.m.is.nagoya-u.ac.jp/>) を使用した。

文 献

[1] 辻孝之, 服部弘, 渡辺正人, “ナイトビジョンシステムの開発,” *Honda R&D Technical Review*, vol.13, no.1, pp.151-158, 2002.

[2] 河上清治, 岩崎克彦, 西田誠, “レーンキーピングアシストの紹介,” *Toyota Technical Review*, vol.55, no.1, pp.54-59, 2006.

[3] 島伸和, 馬場崎正博, 秋月義樹, 本田加奈子, 樋口崇, 東田博文, “運転支援システム用フュージョンセンサの開発,” *富士通テック情報*, vol.19, no.1, pp.31-39, July 2001.

[4] D.M. Gavrila, “Pedestrian detection from a moving vehicle,” *Proc. European Conference on Computer Vision*, pp.37-49, 2000.

[5] 望月大介, 矢野良和, 橋山智訓, 大熊繁, “車載カメラによる歩行者検出のための背景除去及びアクティブ探索を用いたテンプレートマッチングの高速化,” *電子情報通信学会論文誌*, vol.J87-D-II, no.5, pp.1094-1103, May 2004.

[6] 北村康弘, 畑山俊輝, “自動車運転時の聴覚情報への対応課題がドライバーに与える影響 —生理的・行動的变化から—,” *国際交通安全学会誌*, vol.30, no.3, pp.327-332, Sept. 2005.

[7] 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.

[8] 田久保宣晃, “交通事故データによる運転者のヒューマンエラーと心的負荷の一考察,” *国際交通安全学会誌*, vol.30, no.3, pp.299-308, Sept. 2005.

[9] 堀口研一, 熊野史朗, 山口大助, 佐藤洋一, 須田義大, 鈴木高宏, “ドライバの頭部姿勢及び自転車情報を用いた脳見状態推定手法,” *自動車技術会 2007 年秋季学術講演集*, no.100-07, pp.1-6, Oct. 2007.

[10] 秋山知範, 稲垣敏之, 古川宏, 伊藤誠, “視点移動の解析によるドライバの漫然状態検出,” *ヒューマンインターフェースシンポジウム 2005 論文集*, pp.345-350, Sept. 2005.

[11] J.M. Wolfe, “Guided Search 4.0: Current progress with a model of visual search,” in *Integrated Models of Cognitive Systems*, ed. W.D. Gray, pp.99-119, Oxford University Press, 2007.

[12] A. Treisman, and S. Gormican, “Feature analysis in early vision: Evidence from search asymmetries,” *Psychological Review*, vol.95, no.1, pp.15-48, Jan. 1988.

[13] 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, Nov. 1998.

[14] V. Navalpakkam, and L. Itti, “Modeling the influence of task on attention,” *Vision Research*, vol.45, no.2, pp.205-231, Jan. 2005.

[15] G. Paulmier, C. Brusque, V. Carta, and V. Nguyen, “The influence of visual complexity on the detection of targets investigated by computer generated images,” *Lighting Research and Technology*, vol.33, no.3, pp.197-205, Sept. 2001.

[16] 木村文香, 目加田慶人, 高橋友和, 井手一郎, 村瀬洋, 玉津幸政, “運転者支援のための交通信号機視認性数値化手法,” *電気学会論文誌 C*, vol.130, no.6, pp.1034-1041, July 2010.

[17] V. Sharma, and J.W. Davis, “Integrating appearance and motion cues for simultaneous detection and segmentation of pedestrians,” *Proc. Conf. on Computer Vision*, pp.1-8, Oct. 2007.