

# 動的な背景の中から人物だけを 検出する画像認識技術

## ステレオカメラで安全運転を支援

人と車両の事故による死亡及び重傷者は毎年1万人を超え、その半数以上は運転者側に過失があると言われています。運転者に周囲の歩行者の存在を知らせることで、接触事故を防止したり、被害を軽減することができます。

自動車の安全運転を支援するため、車載カメラで周囲の歩行者を検出する画像認識技術の開発が進められています。車載カメラの場合には、車の動きに伴って時々刻々変化する背景の中から、動きのある様々な大きさ、姿勢、服装の人物を検出する必要があり、従来の方式では検出性能が不十分でした。

今回東芝は、ステレオカメラを用いて歩行者を高精度に検出する技術を開発しました。これにより、誤検出を従来の1/10以下に低減することができるようになりました。

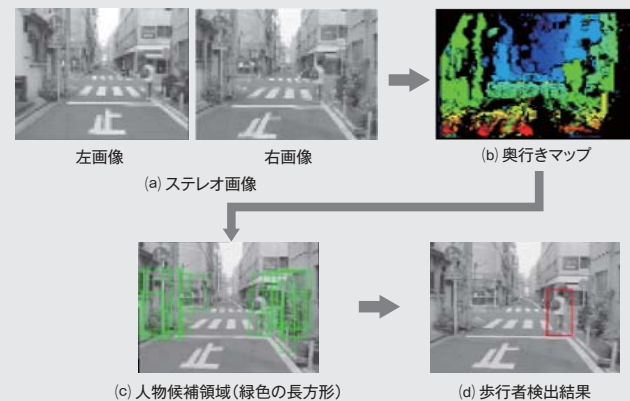


図1. ステレオカメラを用いた歩行者検出処理の概要 — ステレオ画像から各点の奥行きを計算し、路面上の物体から歩行者候補を抽出します。これらの候補からパターン識別によって、歩行者を検出します。

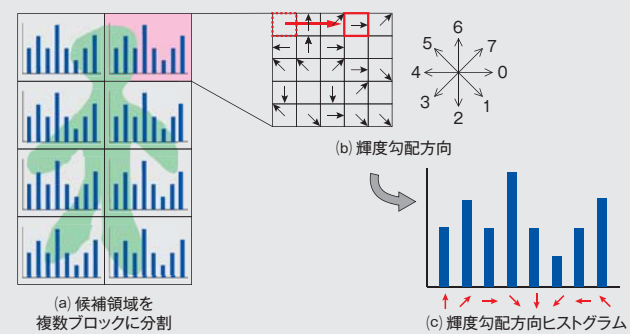


図2. HOG特徴量の概要 — HOG特徴量は、縦・横・斜め方向のエッジがブロック内にどの程度あるかを定量化します。

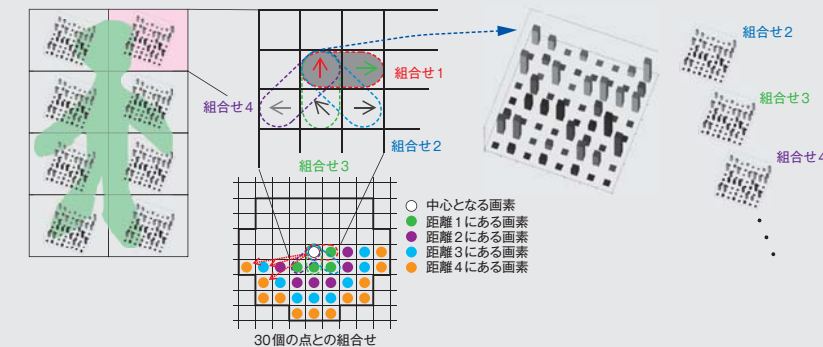


図3. CoHOG特徴量の概要 — CoHOG特徴量は、二つの点で、エッジ方向のどのような組合せがどのくらいの頻度で同時に存在(共起)するかを、4種類の距離にある30個の点との間で調べます。

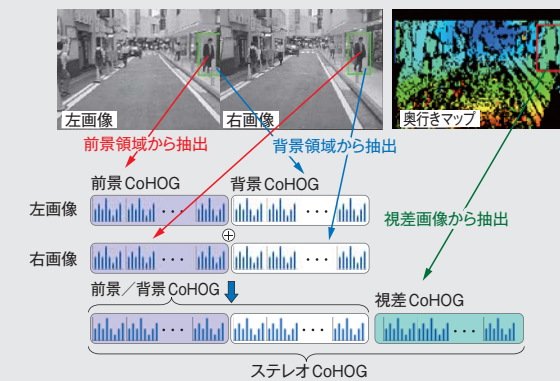


図4. ステレオCoHOG特徴量の概要 — 視差に基づいて前景と背景を分離し、各領域から別々に特徴量を抽出します。左右画像からの特徴量を統合し、それと視差画像から抽出した特徴量を合わせてステレオCoHOGを得ます。

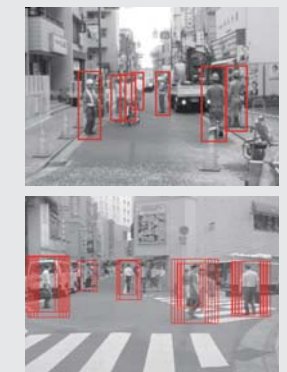


図5. 歩行者検出結果の例 — 歩行者として検出した結果を赤枠で示します。様々なシーンでも精度良く歩行者を検出できます。

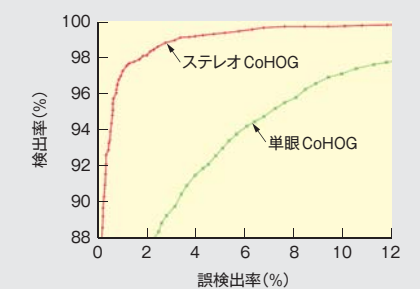


図6. ステレオCoHOGと単眼CoHOGによる識別性能の比較 — 左上に向かうほど識別性能が高いことを意味します。ステレオCoHOGを用いることで、誤検出が1/10以下になります。

### 車載カメラの歩行者検出技術

路上に固定されたカメラで歩行者を検出する場合には、背景が静止しているため、動きのある領域を人物候補として抽出できます。しかし、車載カメラの場合、時々刻々変化する背景の中から、動きのある様々な大きさ、姿勢、服装の人物を検出する必要があります。

東芝は、こうした難しい課題を解決するため、2台のカメラを用いた、ステレオ視に基づく検出方法を開発しています。画像処理の概要を図1に示します。

まず、左右に並べた2台のカメラでステレオ画像(a)を撮影し、画像間で対応する点を求めて、奥行きマップ(b)を作成します。ここでは、近いものほど赤く、遠いものほど青く表示しています。次に、奥行きマップから対象シーンの

3次元構造を推定し、検知範囲内にある、大きさや形状が人間に近い物体を人物候補領域として検出します(c)。ステレオ計測することで物体までの距離がわかるので、適切なサイズの検出枠を設定できます。この段階で、道路領域や、交通標識など人間の身長に比べて極端に縦長の物体、あるいは、横向きのように横幅が広い物体を除外します。更に、これらの人物候補領域を人物とそれ以外のものにパターン識別することで、歩行者を検出します(d)。

ここで、どのような特徴を用いてパターン識別を行うかが重要になります。

### 人物検出のための画像特徴量

人物検出のためによく使われる特徴量として、輝度勾(こう)配方向ヒストグラム(HOG)があります。HOG特徴

量は、候補領域を複数のブロックに分割して、各ブロックの輝度勾配方向をヒストグラム化したものです(図2)。輝度勾配方向とはエッジの方向のことで、HOG特徴量は、縦・横・斜め方向のエッジがブロック内にどの程度あるかを定量化します。そのため、明るさの変化などの影響を受けにくく、形状の変化に対しても頑健であると言われています。

当社は、HOG特徴量を識別性能の面で改良した、輝度勾配方向共起ヒストグラム(CoHOG)を開発しています。HOG特徴量が各点のエッジ方向をヒストグラム化しているのに対して、CoHOG特徴量では、2点のエッジ方向の組合せをヒストグラム化します(図3)。CoHOG特徴量を用いることで、HOG特徴量の長所を継承しつつ、物体の形状をより詳細に識別できます。

### ステレオCoHOG特徴量

しかし、これら従来の特徴量を用いたパターン識別では、例えば背後に複雑な形状のパターンがあるような場合には、人物の特徴が背景に埋もれるため正しく識別できないことがあります。

そこで今回、ステレオ画像を利用して背景の影響を軽減した特徴量として、“ステレオCoHOG”を新たに開発しました(図4)。まず、ステレオ画像間の視差を利用して左画像と右画像の前景と背景を分離し、それらの各領域から別々にCoHOG特徴量(“前景CoHOG”と“背景CoHOG”)を抽出します。次に、左右画像から抽出したこれらの特徴量の各要素を加算し、“前景/背景CoHOG”とします。更に、視差画像からも視差の相対値に基づく特徴量

“視差CoHOG”を抽出します。こうして求めた前景/背景CoHOGと視差CoHOGを連結した特徴量をステレオCoHOGと呼びます。ステレオCoHOG特徴量を用いて、様々なシーンで人物を検出した結果を図5に示します。背景が異なっても、精度良く人物を検出できます。

ステレオCoHOG特徴量の識別性能を、従来のCoHOG(以下、単眼CoHOGと呼ぶ)と比較した結果を、図6に示します。図は、実験車両で市街地を走行しながら撮影した画像データを、二つの特徴量を用いて識別した結果です。縦軸と横軸は、検出率(検出された人数/実際の人数)と誤検出率(まちがった検出の数/総検出数)であり、左上に向かうほど識別が高精度であることを示します。この結果から、ステレオ

CoHOGの識別性能が、単眼CoHOGに比べて高いことがわかります。例えば、検出率96%と比較すると、単眼CoHOGの場合の誤検出率が約8%であるのに対して、ステレオCoHOGでは約0.75%であり、誤検出が1/10以下に低減されています。

### 今後の展望

今後は、今回開発したステレオCoHOG特徴量を用いた画像認識技術の更なる高性能化を目指すとともに、ハードウェア化についても検討を進め、自動車の安全運転を支援していきます。

### 服部 寛

ビジュアルプロダクツ社  
コアテクノロジーセンター  
デジタルプロダクツ技術開発部主務