3T-02

危険要素抽出の高精度化に基づく高安全知能システムの構成

大河原 茂樹† 亀山 充隆†

東北大学大学院情報科学研究科†

1. まえがき

実世界環境で生活する人間の周囲には多様な危険が存在する．そのため，種々の危険要素の抽出に基づき人間にふりかかる危険を事前に検出し警報を与える高品質な高安全知能システムの開発が望まれる．高安全知能システムは，危険要素を抽出し統合するという段階を経て最終的に警報を与えるというプロセスからなる．システムを構築するにあたり，危険要素が存在するときはできるだけ見逃す確率を小さくすることが求められる．本稿では，入力画像に対して複数種類の特徴量により得られた認識結果の論理和や多数決をとることにより危険要素抽出の高精度化を達成する手法を提案し，評価を行う．

2. 危険に対する起因事象の選択

実世界には多様な危険起因事象が多数存在しているため，全ての危険起因事象を考慮しようとすれば，オンライン処理速度に追従できるようなシステムの構築は困難であり，膨大なシステム開発コストも要する．従って，全ての危険起因事象が引き起こすリスクの期待値の大きさによって優先順位を評価し[1]，それに従いシステムがカバーする起因事象をいくつか選択する．

3. 危険要素の統合

高安全知能システムは，危険要素を抽出し統合するという段階を経て最終的に警報を与えるというプロセスから成る．一例として強盗に遭遇した場合の危険検出を行うための危険要素の統合を考える．危険要素として以下を定義する．

a:恐怖を感じている表情をしている人が存在する．
b:ナイフを持っている人物が存在する．
c:ピストルを持っている人物が存在する．
d:その他の危害を与える恐れのある物体を持っている人物が存在する．
e:顔を隠している人物である．

とすれば，強盗犯がいるという必要条件は次式のように表わされる．

$$A \rightarrow a \wedge (b \vee c \vee d \vee e)$$

対偶をとると

$$\overline{a} \vee (\overline{b} \wedge \overline{c} \wedge \overline{d} \wedge \overline{e}) \rightarrow \overline{A}$$

となり，強盗ではないという確実な条件を与えることができる．これに基づき，警報を発しない条件という観点からシステムを構成する．

Design of a Highly-Safe Intelligent System Based on Accurate Extraction of Danger Components

† Shigeki Ookawara , Michitaka Kameyama
† Graduate School of Information Sciences . TOHOKU University

4. 危険要素の抽出

本稿では，画像情報を用いて危険要素の抽出を行う．図１に示すように，１つの特徴量を用いて入力画像から特徴抽出を行い，画像認識の一手法であるBag-of-Featuresに基づき[2]，サポートベクターマシン（SVM）を用いて入力画像の識別を行う．図２に示すように，学習用画像群から特徴抽出を行い，代表的な局所パターンであるVisual wordsに基づき，モデル用画像群，未知画像から画像毎のヒストグラムをそれぞれ作成する．図 3 に示すように，SVM にモデルヒストグラムを学習させ，判別モデルに基づき未知画像ヒストグラムの識別を行うことで，危険要素の抽出を行う

図１：危険要素抽出モデル

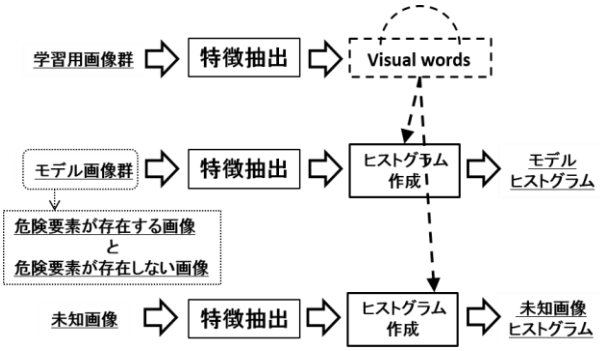

図２：Bag-of-Features を用いたヒストグラム作成

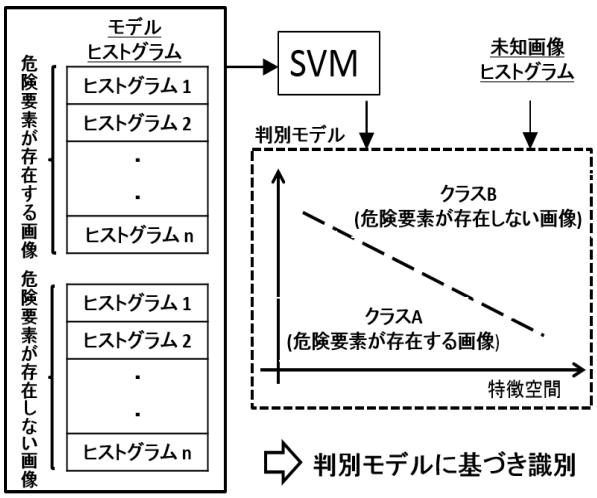

図３：SVM を用いた危険要素抽出

Copyright ©2015 Information Processing Society of Japan.
All Rights Reserved.

今回，特徴量の違いによる認識率を調べるために，例として５つの特徴量 SIFT，DSIFT，PHOW，LBP，HOG を用いて[3]-[5]，危険要素の一例であるナイフの認識を行った．ナイフが存在する入力画像 100 枚，ナイフが存在しない入力画像 100 枚を用いた場合の危険要素の認識結果を表１に示す．

表１：５種類の特徴量に対する認識率

特徴量	ナイフが存在する画像		ナイフが存在しない画像	
	正しい認識	誤った認識	正しい認識	誤った認識
SIFT	73%	27%	74%	26%
DSIFT	74%	26%	92%	8%
PHOW	79%	21%	93%	7%
LBP	71%	29%	67%	33%
HOG	65%	35%	65%	35%

ここで，同じ入力画像でも，特徴量の違いによって認識が正しく行えたり行えなかったりすることに注意をする必要がある．

5. 危険要素抽出の高精度化

危険要素が存在するときはできるだけ見逃す確率を小さくするために，図４に示すように，入力画像に対して，複数種類の特徴量により得られた認識結果（正しい認識が行われた場合“1”，その他“0”）の論理和や多数決による高精度の手法を提案する．

提案手法の評価のために，ナイフが存在する入力画像 100 枚，ナイフが存在しない入力画像 100 枚を対象に，５つの特徴量 SIFT，DSIFT，PHOW，LBP，HOG を用いた場合に，複数の特徴量を用いた認識結果に対して，それらの論理和や多数決をとることにより，危険要素抽出の高精度化を行った．２種類の特徴量により得られた認識結果の論理和をとった場合の認識結果を表２に，３種類の特徴量により得られた認識結果の論理和，さらに，多数決をとった場合の認識結果を表３に示す．

表２，表３に示すように，ナイフが存在する画像に対して，認識結果を組み合わせた場合，正しく認識できる確率が大きくなっていることがわかる．このように，提案手法では，危険要素Ｘが存在する場合に見逃す確率を小さくすることが可能である．しかし，危険要素Ｘが存在しない場合に誤った認識をする確率が大きくなってしまうという問題が存在する．従って，危険要素Ｘが存在しない場合の誤認識率の許容値を設定し，誤認識率が許容値以下という制約条件を満たしつつ，危険要素Ｘが存在する場合の認識率が最大となるような，特徴量の選択を行う必要がある．

例えば，表２，表３において，ナイフが存在しない画像に対して誤った認識してしまう確率が 15%以下という制約条件を設定した場合，その制約条件を満たしつつ，ナイフが存在する画像に対して正しく認識できる確率が最大となるのは，特徴量 DSIFT と PHOW により得られた認識結果の論理和をとったときである．

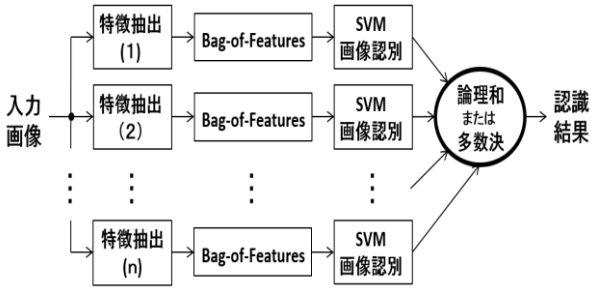

図４：危険要素抽出高精度化モデル

表２：OR による認識結果を採用した場合の認識率

特徴量		ナイフが存在する画像		ナイフが存在しない画像	
		正しい認識	誤った認識	正しい認識	誤った認識
		論理和	論理和	論理和	論理和
SIFT	DSIFT	87%	13%	58%	42%
SIFT	PHOW	89%	11%	60%	40%
SIFT	LBP	89%	11%	44%	56%
SIFT	HOG	85%	15%	48%	52%
DSIFT	PHOW	84%	16%	87%	13%
DSIFT	LBP	88%	12%	63%	37%
DSIFT	HOG	84%	16%	62%	38%
PHOW	LBP	90%	10%	62%	38%
PHOW	HOG	87%	13%	62%	38%
LBP	HOG	78%	22%	52%	48%

表３：多数決による認識結果を採用した場合の認識率

特徴量			ナイフが存在する画像				ナイフが存在しない画像			
			正しい認識		誤った認識		正しい認識		誤った認識	
			論理和	多数決	論理和	多数決	論理和	多数決	論理和	多数決
SIFT	DSIFT	PHOW	91%	78%	9%	22%	57%	91%	43%	9%
SIFT	DSIFT	LBP	95%	74%	5%	26%	42%	81%	58%	19%
SIFT	DSIFT	HOG	94%	70%	6%	30%	46%	79%	54%	21%
SIFT	PHOW	LBP	94%	80%	6%	20%	42%	82%	58%	18%
SIFT	PHOW	HOG	95%	75%	5%	25%	46%	82%	54%	18%
SIFT	LBP	HOG	92%	73%	8%	27%	40%	63%	60%	37%
DSIFT	PHOW	LBP	94%	74%	6%	26%	60%	92%	40%	8%
DSIFT	PHOW	HOG	93%	73%	7%	27%	60%	92%	40%	8%
DSIFT	LBP	HOG	90%	73%	10%	27%	48%	77%	52%	23%
PHOW	LBP	HOG	92%	76%	8%	24%	48%	77%	52%	23%

6. むすび

危険要素が存在するときは，できるだけ見逃しを少なくする危険要素抽出手法は高安全知能システムの構成に有用となる．今後は，この構成方法に基づき，実世界環境で動作できるシステムを構成し，その評価を行う．

参考文献

[1]亀山充隆,樋口龍雄,“知的フォールトトレランス”,電子情報通信会誌,Vol.73,No.11,pp.1233 -1236(1990).
[2]Csurka,G.,Dance,C.R.,Fan,L.,Willamowski,J.,and Bray,C.,“Visual Categorization with Bags of Keypoints”,ECCV International Workshop on Statistical Leaning in Computer Vision pp.59-74(2004).
[3] D. G. Lowe, “Object recognition from local scale invariant features”, Proc. of IEEE International Conference on Computer Vision (ICCV), pp.1150-1157(1999).
[4] N. Dalal, B. Triggs, “Histograms of oriented gradients for human detection”, Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 886-893(2005).
[5] Timo Ojala, Matti Pietik¨ainen, and David Harwood. A comparative study of texture measures with classification based on featured distributions. Pattern Recognition,Vol. 29, No. 1, pp. 51-59(1996).

Copyright ©2015 Information Processing Society of Japan.
All Rights Reserved.