The Minimum Risk Angle for Automatic Target Recognition using the Intersecting Cortical Model

Nils Zetterlund
Royal Institute of Technology
Department of Physics
AlbaNova University Center
S-10691 Stockholm
Sweden
nize1793@student.uu.se

Thomas Lindblad
Royal Institute of Technology
Department of Physics
AlbaNova University Center
S-10691 Stockholm
Sweden
lindblad@particle.kth.se

Ulf Ekblad
Royal Institute of Technology
Department of Physics
AlbaNova University Center
S-10691 Stockholm
Sweden
ulf.ekblad@foi.se

Abstract - *While kids easily find 3-D objects like animals in a scene (e.g. a photograph), this is still not the case for algorithms running on von Neumann computers or neural network chips. The present investigation has two goals: Finding "signatures" of the object in the scene and trying to find out at which observation angle the chance of correct identification is the best. By signatures we mean a vector of reasonable size (say 50 elements). Clearly a cow looks different from the back or from the side. A car is probably more easily identified viewed from the front than it is from above. For a plane it may be the other way around. Thus if we define a general but compact "signature" of the object, it will surely depend on the viewing angle. The problem of finding the most optimal viewing angle is dealt with in this paper.*

Keywords: Image signature, viewing angle, target angle, feature extraction, conditional density estimation.

1 Introduction

Automatic target recognition (ATR) has a wide range of applications, both military and civilian. In many of those a recognition mistake can be very expensive. One way of increasing the performance of an ATR system is to find and use an optimal viewing angle. In this study the minimum risk angle will be defined and several approaches to calculate it will be suggested. The algorithms will be applied on a specific ATR system in a few experiments. This ATR system uses the intersecting cortical model (ICM) to extract feature patterns, so-called signatures, from the images depicting the targets. We will start by making the reader familiar with the ICM.

2 The ICM and Some Definitions

2.1 The ICM and its signature

The Pulse Coupled Neural Network, or the PCNN, is based on findings of the visual cortex system of small mammals [1-4]. The PCNN and its use in image processing have been discussed by Lindblad and Kinser [5] as well as by several investigators applying PCNNs to various fields [6].

The ICM, introduced by Kinser [7], is a simplified version of the PCNN. The input stimulus $\mathbf{S}$ is a digital image in which the pixels S_{ij} are scaled ranging from zero to unity, and fed to a matrix of neurons having states F_{ij} and thresholds Θ_{ij}. They are computed from Eqs. (1-3):

$$F_{ij}[n+1] = fF_{ij}[n] + S_{ij} + \mathrm{W}_{ij}\{Y[n]\} \tag{1}$$

$$Y_{ij}[n+1] = \begin{cases} 1, & if \quad F_{ij}[n+1] > \Theta_{ij}[n] \\ 0, & otherwise \end{cases} \tag{2}$$

$$\Theta_{ij}[n+1] = g\Theta_{ij}[n] + hY_{ij}[n+1] \tag{3}$$

where f, g, and h are scalars and W_{ij} are the connection functions for "communication" between the neurons. The neurons will "fire" and produce binary images $\mathbf{Y}[n]$ having the elements $Y_{ij}[n]$ at each iteration n. These outputs are given by Eq. (2).

Integrating all the pixel values in an output image yields a scalar. By performing this calculation for a series of D output images a vector $\tilde{\mathbf{x}}$ (of dimension D) can be constructed [8]. This is the ICM signature, also called the time signal.

We wanted to test the ICM in order to investigate its usefulness in producing signatures. It has been shown that the ICM signature possesses three important properties, namely that it is invariant, within a certain range of values, to rotation, scaling, and translation [9]. A drawback of the ICM signature is, however, that the generation of it requires many computations. This may be a problem for time critical applications such as some ATR systems. The experimental results of this study will show that this does not have to be a problem.

2.2 Viewpoint and viewing angle

We will here use the ordinary cylindrical coordinates for the scene coordinate system (Fig. 1). The origin is fixed at the focus point $\mathbf{f}$ and the position of the camera is the viewpoint $\mathbf{v}$. The parameters ρ and z are fixed. This implies that the viewpoint $\mathbf{v}$ can be defined unambiguously by the parameter θ which will be referred to as the viewing angle.

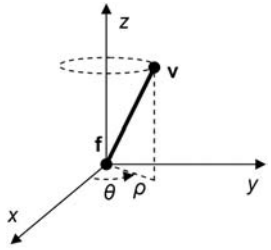

Fig. 1. The scene coordinate system showing the focus point **f**, the viewpoint **v** and the viewing angle θ.

2.3 The target angle

For each target an internal coordinate system, fixed in the target and consisting of three orthogonal axes *x'*, *y'* and *z'* has to be defined. The origin of the target coordinate system is set equal to the origin of the scene coordinate system. The *z*-axis and the *z'*-axis coincide. Call the angle between the *x*-axis and the *x'*-axis the target angle ψ.

2.4 The projection

The photo image will be called the projection **P** of the scene. It is a form of representation of the scene and thus of the target in the scene.

3 ATR using ICM signatures

3.1 System outline

The ATR system considered in this study can be represented as in Fig. 2. A scene containing the target is captured to an image **P** for a viewing angle θ having a known probability density $\mathrm{p}(\theta)$. The image **P** is transformed to an ICM signature $\tilde{\mathbf{x}}$ as described in Sec. 2.1. Subsequently, a feature pattern is extracted from the ICM signature, which in turn is transformed to a decision C_k. The parameters in the *feature extractor/selector* and the *correlator/identifier* are specified using design data. The design data is a set of N numbers of signatures $\tilde{\mathbf{x}}^n$ of known classes C_k^n measured at the viewing angles θ^n. Once the system is designed, the class membership C_k for an unknown image **P** may be estimated.

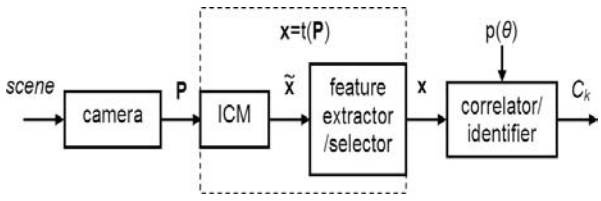

Fig. 2. ATR system outline. The target is captured to an image and classified in a sequence of transformations given the viewing angle density function.

3.2 Feature extraction

Principal component analysis (PCA) is applied to compute the principal component coefficients $x_1^n,\ldots,x_D^n$ so that

$$\mathrm{r}(x_i,x_j)=\sum_{n=1}^{N}\sum_{m=1}^{N}\left(x_i^n-\hat{\mu}_i\right)\left(x_j^m-\hat{\mu}_j\right)=0 \text{ for } i\neq j \qquad (4)$$

where $\hat{\mu}_i=(1/N)\sum_n x_i^n$. The principal component coefficients are sorted and selected so that

$$|\mathrm{r}(x_1,C_k)|\geq\cdots\geq|\mathrm{r}(x_d,C_k)|\geq\xi\max_i|\mathrm{r}(x_i,C_k)| \qquad (5)$$

where $0<\xi\leq 1$. The extracted feature patterns are $\mathbf{x}^n=\left[x_1^n,\ldots,x_d^n\right]^{\mathrm{T}}$ and $d\leq D$.

3.3 Density estimation

The signature **x** can be seen as a measurement on a stochastic variable. The stochastic variable is continuous and dependent on the viewing angle. It is defined by its probability distribution, which can be expressed as a density function $\mathrm{p}(\mathbf{x}\mid C_k,\theta)$. The density functions can be modelled and estimated using the design data. The problem of conditional probability density estimation is not trivial. By studying the relation between the signature **x** and the viewing angle θ some useful conclusions can be drawn (see Sec. 4).

An alternative approach is to model separate density functions $\mathrm{p}(\mathbf{x}\mid C_k)$ for separate intervals of the viewing angle. In such cases, or if the system is designed for a fixed and known viewing angle, the density functions can, for example, be modelled by ordinary (or mixtures of) multivariate normal distributions, multi-layer perceptrons or using the K-nearest-neighbours method [8].

4 Signature functions

Express the projection **P** of a scene S as a function

$$\mathbf{P}=\mathrm{f}(\theta,\psi) \qquad (6)$$

of the viewing angle θ and the target angle ψ. For a non-static scene this function can be written as the sum

$$\mathrm{f}(\theta,\psi)=\mathrm{f}_{\mathrm{det}}(\theta,\psi)+\mathrm{f}_{\mathrm{stoc}}(\theta,\psi) \qquad (7)$$

of a deterministic function $\mathrm{f}_{\mathrm{det}}(\theta,\psi)$ and a stochastic noise $\mathrm{f}_{\mathrm{stoc}}(\theta,\psi)$. Both of these components are strongly dependent on the scene and it is hard to make any general statements on their functional forms.

Suppose that the viewing angle θ is a constant θ_0 and write the projection **P** as a function $\mathrm{g}(\psi)=\mathrm{f}(\theta_0,\psi)$ of the target angle. Viewing an object from one angle and then rotating it 360 degrees naturally gives the same view. In other words the projection function is periodical with the period $T=360/n$, where n is a positive integer determined by the geometry of the target. Generally $n=1$, but for some objects n is larger. For example, a cube with

identical sides may give n equal to four (depending on scene characteristics and how the cube coordinate system is defined).

Since the transformation $\mathbf{x} = \mathrm{t}(\mathbf{P})$ (see Fig. 2) is a mathematical function and the only independent variable is the projection $\mathbf{P}$, the signature function, which we define as $\mathrm{s}(\psi) = \mathrm{t}[\mathrm{g}(\psi)]$, has a period equal to the period of the projection function $\mathrm{g}(\psi)$. The result is

$$\mathrm{s}(\psi) = \mathrm{s}(\psi + T). \tag{8}$$

Given that the signature $\mathbf{x} = \mathrm{s}(\psi)$ is a realization of a periodic stochastic variable it follows that the conditional probability density $\mathrm{p}(\mathbf{x} \mid C_k, \psi)$ is periodic with the same period T. A promising method for estimation of conditional probability densities for periodic variables is introduced in [10] and tested to some extent in [8].

Another property of the signature function is the symmetry. Since the ICM signature $\mathbf{x} = \mathrm{t}(\mathbf{P})$ is invariant to a limited range of rotation of the image [5], the signature of a mirrored image $\mathbf{P}_\mathrm{M}$ and the signature of the original image $\mathbf{P}$ are approximately equal, $\mathrm{t}(\mathbf{P}_\mathrm{M}) \approx \mathrm{t}(\mathbf{P})$. This means that the signature function of a symmetrical target can be defined so that

$$\mathrm{s}(\psi) \approx \mathrm{s}(-\psi). \tag{9}$$

5 Minimum risk angle

5.1 Definition

A reasonable definition of the optimal viewing angle is the viewing angle for which the misclassification probability P(error) of the ATR system is minimized. However, the goal for the ATR system could be another than assigning the pattern $\mathbf{x}$ to the most probable class C_k. Therefore it is common, in the design of pattern classifiers, to introduce costs L_{kj} specifying the penalty associated with assigning a pattern to class C_j when it in fact belongs to C_k. These costs are preset by the user.

The expected cost, also called the *risk*, is a good performance measure of the system. Writing the risk R as a function $\mathrm{R}(\theta)$ of the viewing angle, we can define the minimum risk angle as

$$\begin{aligned} \theta_\mathrm{MR} &= \arg\min_\theta \mathrm{R}(\theta) \\ &= \arg\min_\theta \sum_{k=1}^{c} \int_{\Re_j} \left[\sum_{j=1}^{c} L_{kj}\, \mathrm{p}(\mathbf{x} \mid C_k, \theta) \mathrm{P}(C_k) \right] d\mathbf{x} \end{aligned} \tag{9}$$

where patterns $\mathbf{x}$ falling in region $\Re_j$ are assigned to class C_j and c is the number of classes. By finding and using the minimum risk angle θ_MR, we have optimized the performance of the ATR system. In the following section, we will discuss both which factors are determining the minimum risk angle and how to find it.

5.2 Important factors

The minimum risk angle, defined in Eq. (9), is determined by three factors:

1. the density estimates $\mathrm{p}(\mathbf{x} \mid C_k, \theta)$
2. the costs L_{kj}
3. the prior probabilities $\mathrm{P}(C_k)$

The densities affect the risk by making it higher if the densities of the target classes are overlapping. By overlapping densities we mean that the functions $\mathrm{p}(\mathbf{x} \mid C_k, \theta)$ for different target classes C_k are large for the same values of $\mathbf{x}$. Overlapping densities can have several reasons: the densities are not well estimated, the target classes to be separated look very similar, or the variations within the target classes are large.

The density estimation error comes down to selecting an appropriate density estimation method and having a sufficient amount of design data. The similarity between two target classes, and thereby their densities $\mathrm{p}(\mathbf{x} \mid C_k, \theta)$, may vary strongly with the viewing angle θ. Suppose for example that we are to distinguish a hostile aircraft among a set of friendly ones. Probably the most distinctive view is from under or above, rather than from the side. This may be compared to recognition of autos on a motorway. The auto densities are likely to be overlapping when they are viewed from above and well-separated viewed from the front or from the side. It shall be remembered, however, that the similarity we experience viewing two projections may differ from what is similar in the ICM feature space.

The signature noise within the target classes may also be dependent on the viewing angle. A large noise makes the density functions more flat and smeared out, and can therefore make the densities $\mathrm{p}(\mathbf{x} \mid C_k, \theta)$ more overlapping, leading to increased risk. Consider recognition of persons at an airport. If the camera is directed so that a lot of people and movement are in the scene, the noise will be larger than if the camera direction is chosen so that the projection only shows the person to be recognized in front of a wall.

The costs L_{kj} influence by deciding which target classes that need to be well-separated to make the risk low. For example, assume that the goal is to separate men from women and that we wish to find an optimal viewing angle to do this. Suppose we have access to conditional densities $\mathrm{p}(\mathbf{x} \mid C_k, \theta)$ for a large set of individuals C_k (including both men and women) viewed from different viewing angles θ. The costs L_{kj} can in this case be set to $L_{kj} = 0$ if C_k and C_j are of the same sex, and $L_{kj} = 1$ otherwise. By doing this the minimum risk angle θ_MR will be the viewing angle where the densities $\mathrm{p}(\mathbf{x} \mid C_k, \theta)$ of individuals of opposite sexes are least overlapping. This may mean that the ATR system (using this viewing angle) works very poorly separating individuals within the sexes. An ATR system designed to identify individuals could be obtained by assigning the losses $L_{kj} = 1$ if $j \neq k$ and $L_{kj} = 0$ otherwise. This would probably lead to another

optimal viewing angle θ_{MR} than the one optimal to separate the sexes.

The prior probabilities also influence the minimum risk angle. Consider the problem of separating sexes. Suppose that the ATR system will be used in a home for elderly, where it is very common with old visitors. Suppose further that the set of individuals C_k is of mixed ages a_k. Introducing prior probabilities as $\text{P}(C_k) = \text{P}(a_k)$ and computing the minimum risk angle θ_{MR} would give a different result than if all individuals were equally usual, with prior probabilities $\text{P}(C_k) = 1/c$, where c is the number of individuals. The minimum risk angle would be the angle at which the densities $\text{p}(\mathbf{x} \mid C_k, \theta)$ of the old people are most separated. In the system where the priors are $\text{P}(C_k) = 1/c$, the minimum risk angle might be from the side of the person, while it for the priors $\text{P}(C_k) = \text{P}(a_k)$ could be from above.

5.3 Calculation of the risk

We will now consider how to calculate the risk R for a given viewing angle θ. If the density functions $\text{p}(\mathbf{x} \mid C_k, \theta)$ are modelled as single-kernel Gaussian distributions and if it is only two classes, C_1 and C_2, the risk $\text{R}(\theta)$ can be calculated through integration of the density functions [8]. However, for more complex density models or a higher number of classes c, this approach is difficult. Instead alternative methods to estimate the risk can be applied:

- cross-validation
- Monte-Carlo simulation

In short, cross-validation means that all ICM signatures $\tilde{\mathbf{x}}^n$ in the design data, except for one validation pattern $\tilde{\mathbf{x}}^v$, measured at the given viewing angle θ, are used to specify the parameters in the system. The validation pattern $\tilde{\mathbf{x}}^v$ is then classified by the system to a class C_j. Since the true class membership C_k^v of the validation pattern is known, the cost l_v^θ of the classification is equal to the cost L_{kj}. By repeating this procedure for all V_θ patterns in the design set measured at the given viewing angle θ, the expected cost $\text{R}(\theta)$ can be approximated by the mean cost

$$\hat{\text{R}}(\theta) = \frac{1}{V_\theta} \sum_{v=1}^{V_\theta} l_v^\theta \,. \tag{10}$$

In Monte-Carlo simulations, we start by modelling and estimating the conditional probability densities $\text{p}(\mathbf{x} \mid C_k, \theta)$ using the entire design set. A computer is used to generate patterns $\mathbf{x}$ from the distribution of class C_k with the probability $\text{P}(C_k)$. The generated patterns are classified in the system and since it is known from which class the pattern is generated, the cost of the classification is known. The risk can be estimated in the same manner as in Eq. (10), only that the parameter V_θ is the number of Monte-Carlo simulations at angle θ instead of the number of available patterns in the design set. In the experiments included in Sec. 7, the cross-validation method is used to estimate the risk.

By using any of the discussed methods to estimate the risk, the minimum risk angle can be estimated by

$$\hat{\theta}_{\text{MR}} = \arg \min_\theta \hat{\text{R}}(\theta) \,. \tag{11}$$

The minimization procedure is determined by application conditions. Most central is the sampling-conditions. If it is cheap to sample the scene, the estimated risk $\hat{\text{R}}(\theta)$ may be computed for angles chosen by an optimization algorithm. On the other hand, if it is expensive to sample the scene, it may be more practical to estimate the densities $\text{p}(\mathbf{x} \mid C_k, \theta)$ (on which the estimate $\hat{\text{R}}(\theta)$ relies) isolated from the optimization of Eq. (10).

6 Angle of Maximum Distance

To find the minimum risk angle defined in the previous section many measurements of the targets are required. Images of each target class must be captured many times for each viewing angle to render accurate estimations of the densities $\text{p}(\mathbf{x} \mid C_k, \theta)$ possible. In some applications, such approach is not cost effective. Therefore, a simpler method can be used to give indications on which viewing angles are most likely to give high classifier performances. Such a method will be suggested in this section.

6.1 Definition

Let us consider the problem of separating two target classes, C_1 and C_2. The stochastic real scene can be modelled by a deterministic 3-D scene in a computer. Assume now that the signature estimates, computed using the model, are equal to the corresponding deterministic component of the true scene signatures and, furthermore, that the true stochastic components are normally distributed with means equal to zero and standard deviations independent on the viewing angle. Then it can be proved that the viewing angle, giving the maximum Euclidian distance between the target signatures, is equal to the minimum risk angle defined in the Sec. 4. In practice this means that, through computing the estimates $\hat{\text{s}}_{\text{det}}^k$ as described in Sec. 5.2, the angle of maximum distance

$$\theta_{\text{MD}} = \arg \max_\theta \text{d}\left[\hat{\text{s}}_{\text{det}}^1(\theta), \hat{\text{s}}_{\text{det}}^2(\theta)\right] \tag{12}$$

can be used as an estimation of the minimum risk angle θ_{MR}. The error is small if the computer model represents the most typical (i.e., the expected) scene conditions and if the magnitude of the variations in the scene does not depend so much on the viewing angle. Fig. 3 shows an example of how the distance between two targets (in the signature space) can vary with the viewing angle.

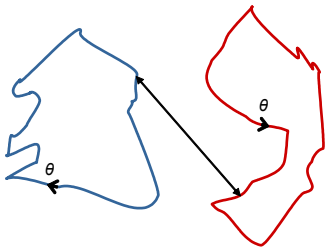

Fig. 3. Schematic illustration of how the Euclidean distance (the length of the line with arrows in both ends) between two targets, C_1 and C_2, varies with the viewing angle θ. The two curves are the signature functions $s^1(\theta)$ and $s^2(\theta)$ plotted over a continuous viewing angle θ. The reason to why the curves are closed is the periodicity of the signature functions. Note that the signatures in this example are two-dimensional.

6.2 Signature predictions

To solve Eq. (12) it may be of interest to predict the signatures for angles that are not sampled. Two methods to predict deterministic signatures will therefore be proposed:

- exact interpolation using linear combinations of radial basis functions
- linear combinations of sinusoids

These methods are chosen because they can be adapted to model smooth, periodical, and symmetrical functions, such as the signature functions described in Sec. 3.3.

Suppose that a target is rotated and the signature of the scene is sampled at different target rotations ψ^n, giving the pairs $(\psi^n, \mathbf{x}^n)$ fulfilling

$$\mathrm{s}(\psi^n) = \mathbf{x}^n \qquad n = 1,\ldots,N\,. \tag{13}$$

Each signature element $x_k = \mathrm{s_k}(\psi)$ can be modelled as a linear combination of radial basis functions

$$\hat{\mathrm{s}}_i(\psi) = \sum_{n=1}^{N} w_{in}\Phi\left(\left\|\psi - \psi^n\right\|\right) \tag{14}$$

where w_{in} is the weight corresponding to the Gaussian base function

$$\Phi(z) = \exp\left(-\frac{z^2}{2\sigma^2}\right) \tag{15}$$

having its centre for the angle ψ^n. The parameter σ controls the smoothness of the base functions. Given a value on the smoothing parameter σ, the weights w_{in} can easily be calculated so that the exact interpolation condition in Eq. (13) is fulfilled [11].

An alternative is to model the functions $\mathrm{s}_i(\psi)$ as a sum of S sinusoids

$$\hat{\mathrm{s}}_i(\psi) = \sum_{s=1}^{S} a_{is}\sin(b_{is}\psi + c_{is}) + \frac{1}{N}\sum_{n=1}^{N} x_i^n \tag{16}$$

where the amplitudes a_{is}, frequencies b_{is}, and phase shifts c_{is} are optimized to minimize the sum of squares error

$$\mathrm{SSE} = \frac{1}{2}\sum_{n=1}^{N}\sum_{i=1}^{d}\left[\hat{\mathrm{s}}_i(\psi^n) - x_i^n\right]^2 \tag{17}$$

where d is the dimensionality of the signature $\mathbf{x}$. Minimizing Eq. (17) does not, for a finite number of terms S, imply that the condition Eq. (13) is fulfilled. Another drawback, in comparison with the radial basis approach, is the computational cost of the optimization. The sum of sinusoids model has, however, an advantage of being somewhat more mathematical intuitive. For example, the periodicity is easily modelled by the condition

$$b_{is} = \frac{360 q_s}{T} \tag{18}$$

where q_s are positive integers and T is the period of the signature function $\mathrm{s}(\psi)$. Methods to choose the number of components S and the integers q_s are yet to be investigated. It is likely that consideration ought to be taken to the sampling frequency and to any information on the spectral contents of the signals $\mathrm{s}_i(\psi)$ that may be available.

By adding the condition $c_{is} \equiv 90$ degrees on the phase shift parameters, all terms are restricted to be cosines. Since cosines are symmetrical, the sum in Eq. (16) is a symmetrical function. In this way signatures can be accurately modelled for symmetrical targets, having signatures fulfilling Eq. (9).

Irrespectively of which method is applied, having the signature estimates, the Euclidean distance can be calculated for any given viewing angle θ. An optimization algorithm can be used to solve Eq. (12) and thereby find the angle of maximum distance θ_{MD}.

7 Experimental results

7.1 Minimum risk angle

Consider the problem of identifying vehicles using a camera placed along side of the road. From which angle should it look to maximize the performance of the ATR system? This question will be answered for a scene modelled in the simulation program Operation Flashpoint™. Experimental evaluation on synthetic images may not give the most realistic results but in order to prove the principle it helps to be able to fully control the data collection. The program includes functionality to simulate changes in daytime and weather (such as overcast and fog). Four vehicle models are used: a

motorcycle, a bicycle, a sports car, and a small car. Each of them is captured on images 12 times for randomized weather and daytime conditions. This is repeated for five viewing angles 270, 290, 310, 330, and 350 degrees, where 270 degrees is from the right side and 360 degrees is from the front. Totally, this gives 240 images. Fig. 4 shows four of them.

Fig. 4. Four vehicles (sports car, small car, motorcycle, and bicycle) seen from four different viewing angles.

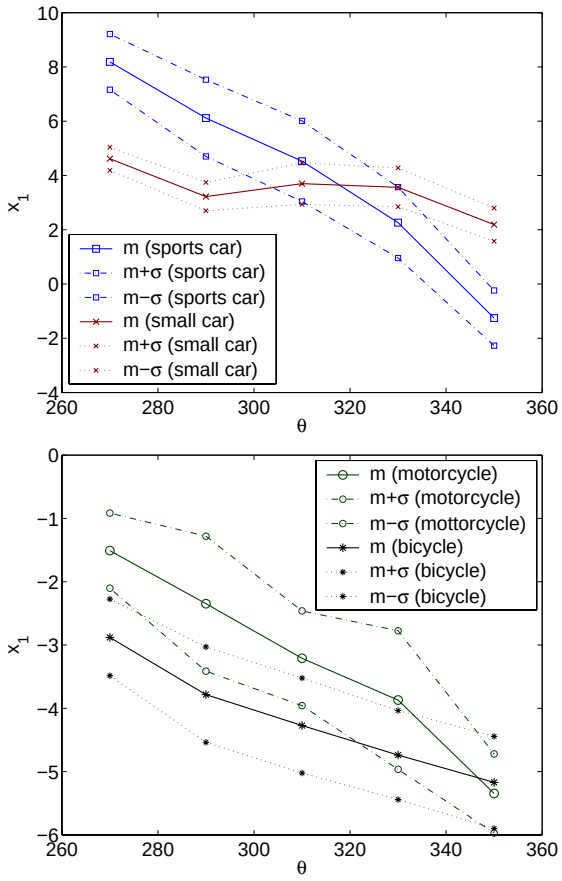

Fig. 5. Means and standard deviations of signature element x_1 for five different viewing angles. By noting that the ranges of the y-axes differ between the two plots, it can be concluded that the vehicles with two wheels (upper plot) are separated from the vehicles with four wheels (lower plot).

The images $\mathbf{P}^n$ are transformed to ICM signatures $\tilde{\mathbf{x}}^n$ of dimensions D=47. Principal component analysis is applied giving the feature patterns $\mathbf{x}^n$. The mean values and the standard deviations of element x_1 of each class are shown in Fig. 5. The densities $p(\mathbf{x}|C_k)$ are estimated separately for each of the five viewing angles. They are modelled as multivariate normal distributions.

The minimum risk angle will be computed for two cases, *case 1* and *case 2*, with two different setups of prior probabilities and misclassification costs. In the first case, *case 1*, all prior probabilities $P(C_k)$ are set to 1/4 and the costs so that all misclassifications are equally expensive. The method described in Sec. 3.2 ($\xi = 0.1$) is used to extract class-correlated features from the ICM signatures. Cross-validation is used to estimate the risks (in this case equal to the misclassification probabilities) presented in Table 1. The low risks indicate that the system works very well for all investigated angles. This can be explained by viewing Fig. 6, which shows how well the classes are separated already in two dimensions. Since a probability cannot be less than zero, we can conclude that the viewing angles 270, 290, 310, and 330 degrees are global minimizers of the estimated risk.

Table 1: Estimated risks R given viewing angle θ.

θ	R (*case 1*)	R (*case 2*)
270	0.0 %	1.0 $
290	0.0 %	1.0 $
310	0.0 %	1.0 $
330	0.0 %	37.8 $
350	2.1 %	1.0 $

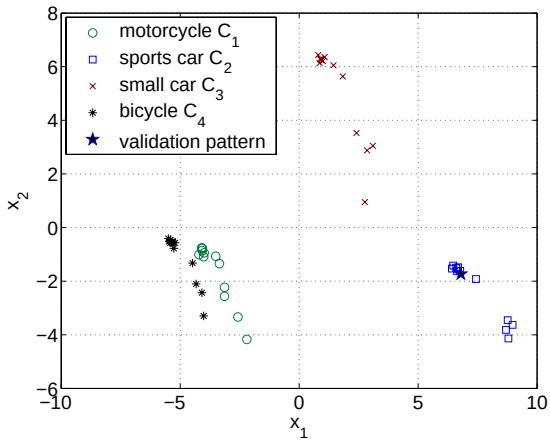

Fig. 6. Validation cycle 1 of 48 for the viewing angle 270 degrees. Note that the validation pattern marked in the plot belongs to class C_2.

The goal of the ATR system may, however, be another than assigning the pattern to the most probable class. For example, suppose that we are searching for the motorcycle C_1, and that we will perform manual controls of all vehicles the ATR system identifies as motorcycles. Suppose that the user would experience a cost of 5000 $ if

the motorcycle passed undetected, but a cost of 100 $ per manual control. This gives the cost matrix

$$\mathbf{L} = \begin{bmatrix} 100 & 5000 & 5000 & 5000 \\ 100 & 0 & 0 & 0 \\ 100 & 0 & 0 & 0 \\ 100 & 0 & 0 & 0 \end{bmatrix} \text{ \$.} \qquad (19)$$

Assume furthermore that the prior probabilities are changed so that $P(C_1)=0.01$, $P(C_2)=0.2$, $P(C_3)=0.3$, and $P(C_4)=0.49$. In other words, the motorcycle C_1 is very rare. Call this problem *case 2*. That the system in the previous example, *case 1*, had optimal performance for several viewing angles indicates that all 47 ICM iterations (or measurements) may not be required. So this time only the three first elements $\tilde{\mathbf{x}}=[\tilde{x}_1,\tilde{x}_2,\tilde{x}_3]^T$ are measured, decreasing the computational load of each decision more than 15 times. The expected costs are estimated using cross-validation (see Table 1). Without the access to the ATR system, the cheapest solution would be to not make any manual controls at all, letting the motorcycles pass undetected. This gives an expected cost of 50 $ per passing vehicle. This means, that by using the optimal viewing angle, the ATR system decreases the cost per decision 50 times. If the non-optimal viewing angle 330 degrees is used, the risk would be reduced much less. Another interesting result is that the optimal viewing angles solving this problem differ from the optimal viewing angles solving the problem in *case 1*.

7.2 Angle of maximum distance

The angle of maximum distance defined in Eq. (12) will be estimated for the Volkswagens shown in Fig. 7. Both of these 3-D models are sampled in front of a homogeneous black background each five degrees.

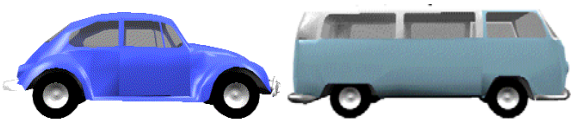

Fig. 7. Digital models of a Volkswagen beetle and a Volkswagen bus.

To evaluate the RBF prediction method discussed in Sec. 5.2, only a part of these data (one sample each 30 degrees) is used to estimate the signature functions, and all of the 72 data points are used for validation. A resulting RBF estimate is shown in Fig. 8. The smoothing parameter is set to 20 and validation gives a root mean square error equal to 27.6×10^{-3}.

The RBF estimates of the two target signatures are used to estimate the distance between the signatures as function of viewing angle. Fig. 9 shows the result. Suppose that we only had access to the estimated distance. Then the angle of maximum distance would be approximated by 90 degrees (according to the maximum value of the estimated curve in Fig. 9). This angle corresponds to the car projections as they are shown in Fig. 7.

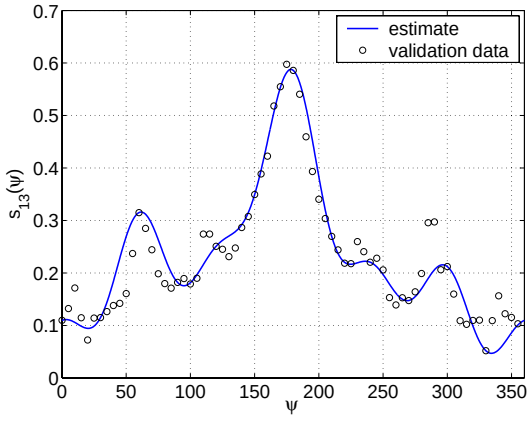

Fig. 8. RBF estimate of signature element $s_{13}(\psi)$ as a function of the target angle.

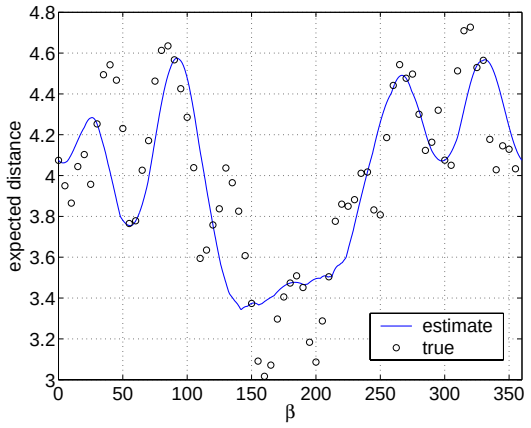

Fig. 9. The distance between the beetle signature and the bus signature given the common target angle β. The estimate is based on RBF interpolations of the beetle signature and the bus signature. The interpolation points are at β=0,30,...,330.

8 Discussion

Several conclusions can be drawn from the simulations estimating the risk. Multivariate normal distributions may be fully adequate to model densities for the noisy ICM signature. The experiments show that optimal performance may be achieved computing no more than three ICM iterations (giving a low computational cost per decision). Moreover, it is proven that the performance of the ATR indeed depend on the viewing angle. It is also shown that the optimal viewing angle changes if the costs or the prior probabilities are changed.

To have a general approach for modelling conditional densities $p(\mathbf{x}|C_k,\theta)$ defined over a continuous viewing angle θ estimated using measurements $\{\mathbf{x}^n,C_k^n,\theta^n : n=1,...,N\}$ would be very valuable in the search for optimal viewing angles and in applications for

which the viewing angle can not be fully controlled (but through the density function $p(\theta)$). Future work will tell if the estimation methods introduced in [8,10] are practical in the context of modelling signature probability densities.

The experiments indicate that the radial basis function interpolation works well for signature prediction. In the future it will be compared to the sum of sinusoids model.

It is important to remember that the proposed methods (minimizing the risk and estimating the optimal viewing angles) can be applied to ATR systems using other feature patterns than the ICM signature. Moreover, the dependent variable does not have to be the viewing angle θ nor the target angle ψ. It could be magnitudes such as light intensity, distance/zoom d to target among many others. Note, however, that with the change of dependent variable the characteristics of the signature function changes concerning periodicity and so on. The signature function $s(d)$ as a function of distance d to the target may have interesting properties since the ICM signature is invariant to (a limited extent of) scaling.

The signature function in this study is a function of a one-dimensional angle. This is not difficult to generalize to a three-dimensional angle (the rotations around all three axes in the scene coordinate system) and can be done for applications in which it is of interest.

Further work will include investigations of what the use of several cameras could entail. The ICM signatures will also be tested on real image data, and comparisons with other methods, such as [12], are on the list for future work. The feature extraction method has flaws which will be amended in future work.

9 Conclusion

It is shown that the ICM signature works well as a feature pattern in the simulated automatic target recognition system. It is also shown that by using an optimal viewing angle, the performance of the system can be increased substantially. The optimal viewing angle depends on several factors, such as scene properties, prior probabilities of the targets, and misclassification costs.

Acknowledgements

The authors would like to acknowledge the discussions with Dr Jason M. Kinser as well as several diploma workers at the department of physics working with the PCNN/ICM projects. These include Jenny Atmer, Martine Gudmundson, and Lars Edvardsson.

References

[1] R. Eckhorn, H. J. Reitboeck, M. Arndt, and P. Dicke. Feature linking via synchronization among distributed assemblies: Simulations of results from cat visual cortex. *Neural Comp*., 2:293-307, 1990.

[2] R. FitzHugh. Impulses and physiological states in theoretical models of nerve membrane. *Biophysics J.,* 1:445-466, 1961.

[3] A. L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology*, 117:500-544, 1952.

[4] J. L. Johnson. Pulse-coupled neural nets: translation, rotation, scale, distortion, and intensity signal invariances for images. *Appl. Opt.,* 33(26):6239-6253, 1994.

[5] Thomas Lindblad and Jason M. Kinser. *Image Processing using Pulse-Coupled Neural Networks*, Springer Verlag, London, 1998, ISBN 3-540-76264-7.

[6] Thomas Lindblad, Mary Lou Padgett, and Jason M. Kinser. Virtual Intelligence/Dynamic Neural Networks, SPIE Volume 3728, 1998.

[7] Jason M. Kinser. A simplified pulse-coupled neural network. In *Proceedings SPIE*, volume 2760, No. 3, pages 563-567, 1996.

[8] Nils Zetterlund. The minimum risk angle or intersected cortical model signatures as feature patterns for automatic target recognition. Diploma work, KTH-Physics, Stockholm, Sweden, 2003.

[9] Ulf Ekblad, Jason Kinser, Jenny Atmer, and Nils Zetterlund. The intersecting cortical model in image processing. In *Imaging 2003*, Stockholm, Sweden, 2003. To be publ. in *Nucl. Instr. Meth. A*; available on www.sciencedirect.com, 2004.

[10] C. M. Bishop and C. Legleye. Estimating conditional probability densities for periodic variables. *Advances in Neural Information Processing Systems,* Volume 7. Cambridge MA: MIT Press.

[11] M. J. D. Powell. Radial basis functions for multivariate interpolation: a review. In J. C. Mason and M. G. Cox (Eds.), *Algorithms for Approximation*, pp. 143-167. Oxford: Clarendon Press.

[12] R. Fergus, P. Perona, and A. Zisserman. Object class recognition by unsupervised scale invariant learning. Proceeeings of the IEEE Conference on Computer Vision and Pattern Recognition, Vol. 2, 2003, pp. 264-271.