Aircraft Recognition from Features Extracted from Measured and Simulated Radar Range Profiles

Aircraft Recognition from Features Extracted from Measured and Simulated Radar Range Profiles

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof.mr. P.F. van der Heijden
ten overstaan van een door het college voor promoties ingestelde
commissie, in het openbaar te verdedigen in de Aula der Universiteit
op 15 oktober 2003, te 12.00 uur

door

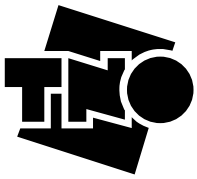
Joris Portegies Zwart

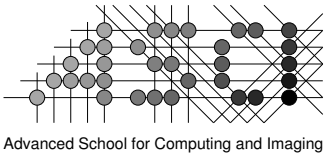
geboren te Amsterdam

Promotiecommissie:

Promotor:	Prof.dr.ir. F.C.A. Groen
Co-promotores:	Dr.ir. B.J.A. Kröse
	Dr.ir. S.J. Gelsema
	Dr. R. van der Heiden
Overige leden:	Prof.dr. P.W. Adriaans
	Prof.dr. S.-E. Hamran
	Prof.dr. P.M.A. Sloot
	Dr. A.R. Webb
	Prof.dr. A.P.M. Zwamborn

Faculteit der Natuurwetenschappen, Wiskunde en Informatica





This work was sponsored by TNO Physics and Electronics Laboratory , The Hague, The Netherlands, and was carried out in graduate school ASCI.

ASCI dissertation series number 92.

Want bepaalde dingen weet je
En dan kun je ze voorspellen

– *Cruijff*, Maarten van Roozendaal

voor mijn ouders

This book was typeset by the author using LaTeX 2ε in combination with the KOMA-Script bundle.

Cover design by Frank Sterk/Edge Pre media.

Contents

1 Introduction

The ability to reliably identify aircraft is an important aspect of air traffic safety. Civilian air traffic controllers need to be constantly updated on the status of aircraft moving through the local airspace. In military scenarios, the need to reliably identify aircraft is even more stringent, since erroneous identification could easily result in friendly fire incidents.

A common technique for identification of military aircraft is Identification Friend Foe (IFF). IFF identification is initiated when the *interrogator* transmits a *challenge* to the aircraft. Friendly aircraft are supposed to be equipped with a *transponder*, which replies to the challenge by transmitting an identification code to the interrogator. Some IFF modes of operation require more information to be included in the reply, such as the current aircraft altitude. Hostile aircraft will in general not be able to respond properly to the challenge because of the lack of a (compatible) transponder, and will therefore be identified as hostile (or at least not friendly). Various other identification techniques are used in combination with IFF. For example, friendly aircraft can be required to limit their flight path to pre-defined regions of airspace called *corridors*.

Civilian aircraft use a technique similar to IFF called Secondary Surveillance Radar (SSR). Although SSR like IFF provides information on aircraft type, its primary purpose is to keep track of the location of civilian aircraft. Until recently it could be safely assumed that all civilian aircraft carried a SSR transponder. However, on September 11th, 2001, terrorists hijacked several civilian aircraft, which were then used to attack both the World Trade Center and the Pentagon, resulting in the loss of over three thousand lives. By disabling the SSR transponders the terrorists prevented air traffic controllers to detect the altered flight path of the aircraft.

The fundamental drawback of techniques like IFF and SSR is that they require active cooperation from friendly aircraft. However, the events of September 11 and many other incidents in recent history show that the cooperation of friendly aircraft is not guaranteed. Friendly aircraft have failed to produce valid IFF replies for a number of reasons, including hardware failure and human error.

Non-Cooperative Target Recognition (NCTR) techniques do not require the active participation of friendly aircraft. Instead, they rely on sensor measurements to independently obtain information on the aircraft. The goal of NCTR is to infer the original aircraft type from these measurements. This is a classic example of a *pattern recognition* problem.

The term pattern recognition describes a wide area of research, which includes topics from the fields of signal processing, computer science, and statistics. Typical pattern recognition problems range from speech recognition and handwritten character classification to robot localisation and fault detection in industrial machinery. The problem of deciding from which aircraft a particular measurement originates is an example of a *classification* problem. In classification problems, the task is to develop an algorithm (or *classifier*) which is capable of deciding to which of a set of distinct classes a given measurement belongs. In the case of aircraft classification the classes are different types of aircraft, and it is the task of the classifier to decide from which aircraft a given measurement originates.

Classifiers can be based on a set of rules derived from expert knowledge of the problem domain, or they can be based on the statistical properties of a *training set*: a collection of measurements for which the corresponding class is known. In the latter case, the resulting *statistical* classifier assigns a new measurement to the class which most likely generated the measurement [1, 2, 3, 4]. *Model-based* classifiers assume some explicit functional form of the

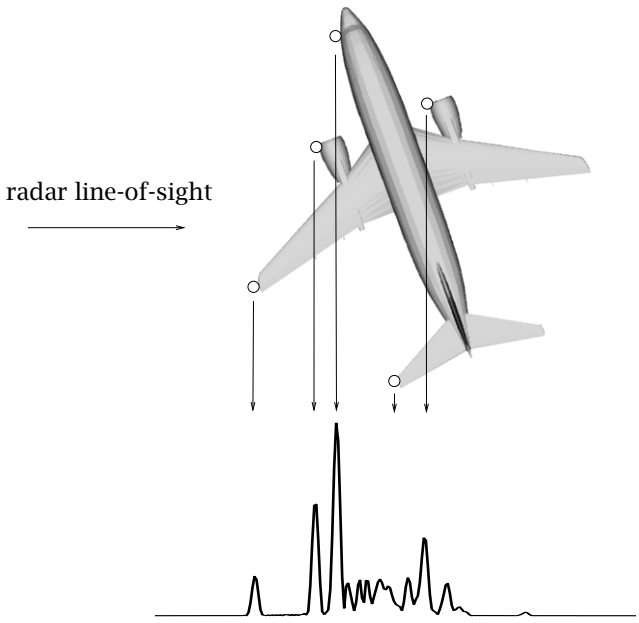


Figure 1.1: Example of a range profile of a Boeing 737-500. The radar is situated on the left hand side. Radar returns from the scatterers on the aircraft (circles) are projected onto the line-of-sight, resulting in a radar range profile (bottom). Taken from [7], with permission.

true distribution of the measurements (for instance, a Gaussian distribution), and infer the optimal model parameters from the training set. Classifiers based on *template matching*, such as the nearest neighbour classifier, implicitly assume some distribution by defining a *similarity measure* between an unknown measurement and the measurements in the training set. Finally, statistical classifiers can also be constructed using *discriminant analysis*, which uses the data set to construct optimal decision boundaries separating the different classes present in the data set.

Many different sensors can be used to obtain measurements for constructing a training set. Radar is an attractive sensor for aircraft classification, which compares well to alternatives such as imaging or infra-red sensors. It is capable of operating over ranges up to several hundred kilometres, operates equally well by day or by night and is insensitive to a wide range of weather conditions. *Radar range profiles* are a specific type of radar measurements which are often used for classifying aircraft [5, 6, 7, 8, 9, 10, 11, 12, 13], but have also been used for classifying ships [14, 15, 16, 17, 18, 19] and ground based vehicles [20]. Radar range profiles are essentially one-dimensional 'images' of aircraft, measured along the line-of-sight between the radar and the aircraft. An example of a range profile is shown in figure 1.1. Range profiles have a number of properties which makes them suitable measurements for classification purposes. They can be measured relatively quickly, and can be measured at any orientation of the aircraft. Furthermore, the requirements on radar hardware are relatively modest.

The use of range profiles for aircraft classification does however have drawbacks, which this thesis aims to solve. The main drawbacks are the need for translation invariant classification, the limited accuracy of estimates of aircraft orientation with respect to the radar, and the difficulty of obtaining sufficient training data.

The need for translation invariant classification results from the fact that changes in the distance between the radar and the aircraft causes cyclic shifts of the resulting range profile. The distance between radar and aircraft is usually not known with high enough precision to

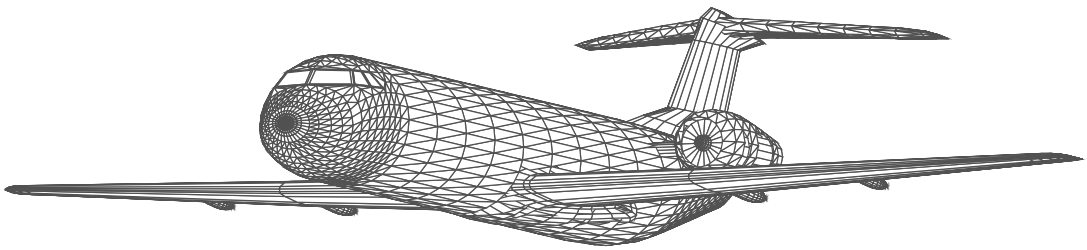


Figure 1.2: A CAD model of a Fokker 100.

correct for this effect, and so a classifier has to be invariant to arbitrary shifts of the range profile(s) to be classified.

Translation invariant classification is often achieved by using a translation invariant similarity measure for comparing range profiles [11, 14, 18, 21]. The main drawback of these approaches is that they are relatively slow. Another solution is to extract translation invariant features from the range profiles [12, 22], which has the drawback that potentially discriminative information is removed from the data. This thesis solves the problem of translation invariant classification by defining a translation invariant representation of range profiles. This method, described in chapter 3, aims to provide fast classification, but does not discard any potentially useful information from the range profiles.

Range profiles are highly dependent on the orientation of the aircraft with respect to the radar, which is usually expressed as an *aspect angle*, defined in chapter 2. In [11] it is shown that the performance of a nearest neighbour classifier increases when aspect angle estimates are used to only compare range profiles with those profiles in the training set which were obtained at comparable aspect angles. Furthermore, the availability of accurate aspect angle information is important for a number of radar signal processing techniques.

Although aspect angles of aircraft can be estimated from tracking data, the estimates suffer from both systematic and random errors in the order of a few degrees [23]. Since range profiles can differ dramatically even over changes in aspect angle smaller than one degree, increasing the accuracy of aspect angle estimates could significantly increase classification performance. Chapter 4 describes a method which combines aspect angle estimates obtained from tracking data with the corresponding radar measurements to improve the accuracy of aspect angle estimates.

The main drawback of using radar range profiles for aircraft classification is that very large training sets are required for constructing reliable statistical classifiers. The reason for this is twofold. First, range profiles are generally represented as high-dimensional vectors containing several hundreds of elements. It is well known that constructing statistical classifiers in high-dimensional vector spaces requires huge amounts of training data [4]. Second, the large dependency of range profiles on aspect angle causes large within-class variability, and so a large training set is required to accurately reflect the statistical distribution of range profiles. Unfortunately, measuring range profiles in a controlled experiment is both expensive and time-consuming. Also, not all aircraft types are available for measuring. Obtaining enough range profile measurements to construct a sufficiently large training set is therefore not a feasible option.

Simulated range profiles provide an alternative source for constructing a training set. Simulated range profiles can be obtained using electro-magnetic simulation software together with CAD models of aircraft such as shown in figure 1.2. Simulated range profiles can be obtained

quickly, cheaply, and at any orientation of the aircraft. Furthermore, CAD models might be obtained even for aircraft not available in real life.

Simulated range profiles are in many ways different from measured range profiles. If a simulated training set is used to construct a classifier for measured range profiles, it should base its decision solely on those features which are present in both simulated and measured range profiles. This observation leads to three research questions answered in chapters 5 and 6: what are the common features present in both simulated and measured range profiles, how can these features be extracted from the data, and how can the extracted features be used for aircraft classification?

The differences between simulated and measured range profiles are often ignored when constructing a classifier from a simulated training set, assuming that simulated and measured profiles can be compared directly. Although reasonable classification performance can be obtained in this way, it does not provide insight in exactly which properties of simulated and measured range profiles provided the basis for comparison. Also, it was shown in [7, 11] that correcting for some of the differences between simulated and measured range profiles results in increased classifier performance.

In this thesis it is argued that the common features between simulated and measured range profiles are the positions of local amplitude maxima, i.e. peaks, in range profiles. Chapter 5 defines a generative model of peak positions, which is capable of generating the statistical distribution of peak positions, given the aspect angle at which the aircraft is observed. The parameters of this model have clear physical interpretations. They include the three-dimensional position of aircraft parts which contribute significantly to the observed range profile (so-called *scatterers*), as well as the probability that a given scatterer is visible at a given aspect angle.

A statistical classifier for measured range profiles, based on the generative model of peak positions, is constructed in chapter 6. This classifier is trained solely from simulated training data, but is capable of classifying measured range profiles. It can be used to classify single range profiles as well as sequences of successive range profiles. Finally, it is translation invariant, and treats the uncertainty in aspect angle estimates in a statistically sound fashion.

Before presenting the actual research on radar range profile processing and classification, this thesis starts by providing an overview of the basic principles of obtaining and pre-processing radar range profiles in chapter 2.

2 Radar Range Profiles

2.1 Introduction

High resolution radar (HRR) range profiles are essentially one-dimensional 'images' of radar targets (see figure 1.1). If a range profile is measured with sufficient resolution, the separate contributions of the *scatterers*, i.e., the parts of the aircraft that strongly reflect the radar energy, are resolved. Range profiles therefore provide information on the geometry of the aircraft, and so they are suitable features for automatic aircraft classification. This chapter summarises the main properties of radar range profiles, after first defining the appropriate coordinate systems in section 2.2.

Radar range profiles are obtained by transmitting a radar signal that comprises a wide range in frequency, using either a single pulse (a chirp), or a series of pulses with increasing frequency. A range profile is defined as the squared magnitude of the coherent complex radar returns, after optional pre-processing. All phase information is usually discarded. The resolution of the range profile is inversely related to the transmitted bandwidth B, as will be shown in section 2.3. However, optional windowing in the Fourier domain usually decreases the resolution (see section 2.5).

Radars usually measure a sequence (or *leg*) of consecutive range profiles over a period of time. The motion of the aircraft during this period causes the aircraft to appear at different positions in different profiles. This effect is known as translation range migration (TRM). Aircraft motion also influences the pose of the aircraft with respect to the radar (or *aspect angle*). Occlusion of scatterers, rotational range migration (RRM) and speckle are effects caused by aircraft rotations which greatly influence range profiles. Section 2.4 discusses these sources of range profile variability in detail.

Throughout this thesis simulated range profiles will be used as the basis for constructing classifiers. Section 2.6 discusses the basics of range profile simulation, and the main differences between measured and simulated range profiles.

Finally, section 2.7 describes the details of the data set used in the experiments performed for the research described in this thesis.

2.2 Coordinate Systems

This section introduces a coordinate system, used throughout this thesis, in which locations relative to the aircraft and aircraft orientation with respect to the radar are defined.

A set of unit vectors $(\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z)$ attached to the aircraft define the axis of the *aircraft-fixed* coordinate system shown in figure 2.1. The positive x-axis points in the direction from the tail to the nose of the aircraft. The positive y-axis points from the wing tip at the starboard side of the aircraft to the wing tip at port side. Finally, the z-axis points from the bottom of the aircraft to the top. Note that the origin of the aircraft-fixed coordinate system is fixed but arbitrary.

Radar line-of-sight is defined as the direction from which the radar illuminates the aircraft. The line-of-sight is expressed in the aircraft-fixed coordinate system as a unit vector $\mathbf{s}$ attached to the origin.

The orientation of the aircraft with respect to the radar is determined by the orientation of $\mathbf{s}$ with respect to the principal axis of the aircraft-fixed coordinate system. Aircraft orientation

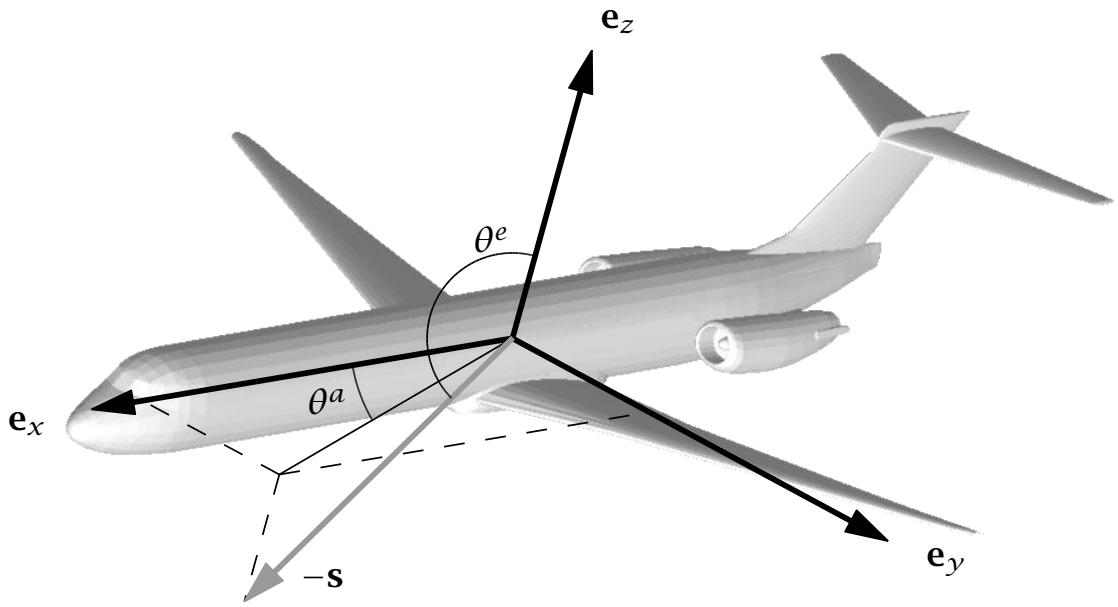


Figure 2.1: The definition of aspect angle $\boldsymbol{\theta}$ in the aircraft-fixed coordinate frame $(\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z)$. The unit vector $-\mathbf{s}$ is the reverse line-of-sight.

is generally described as a two-dimensional *aspect angle* $\boldsymbol{\theta} = (\theta^a, \theta^e)$,[1] Aspect azimuth, θ^a, is defined as

$$\theta^a = \arctan\left(\frac{-\mathbf{s}\cdot\mathbf{e}_y}{-\mathbf{s}\cdot\mathbf{e}_x}\right), \tag{2.1}$$

i.e., θ^a is the angle between the positive x-axis and the projection of $-\mathbf{s}$ on the plane spanned by $\mathbf{e}_x$ and $\mathbf{e}_y$. Aspect elevation, θ^e, is defined as

$$\theta^e = \arccos\left(-\mathbf{s}\cdot\mathbf{e}_z\right), \tag{2.2}$$

i.e., θ^e is the angle between the positive z-axis and $-\mathbf{s}$. Figure 2.1 shows the geometrical construction of both aspect angles.

Aspect azimuth is zero if the aircraft is viewed from nose-on and 180 degrees if viewed from tail-on. If the aircraft is observed from the starboard side then $\theta^a < 0$. If the aircraft is observed from the port side then $\theta^a > 0$. Since aircraft are usually symmetric, it can be assumed that a range profile measured at an aspect azimuth θ^a is identical to a profile measured at $-\theta^a$ (for a given aspect elevation θ^e).

Estimates of the aspect angles of an aircraft in flight can be obtained either from onboard equipment, like INS or GPS-based systems, or they can be derived from *tracking data* – a sequence of three-dimensional aircraft positions, which can be obtained from a tracking radar. While onboard equipment provides the most accurate aspect angle estimates, such equipment is only available in controlled experiments. When performing radar measurements of targets of opportunity, one has to rely on estimates from tracking data.

Aspect angle estimates from tracking data are obtained by combining the tracking data with aerodynamic assumptions on the most likely pose of the aircraft given its flight path. Generally, one assumes that an aircraft moves in the direction of its nose. However, e.g. wind conditions can cause the aircraft to move in a slightly different direction, which results in a systematic error on aspect angle estimates. If information on the weather condition is available, the sys-

[1] Although three angles are required to completely specify aircraft orientation, rotations of the aircraft in the plane perpendicular to the line of sight do not influence range profile measurements. Therefore, a description of aircraft pose in terms of aspect angles is sufficient.

tematic error on aspect angle estimates can be considerably reduced [23].

Aspect angle estimates also suffer from random errors which are the result of small rotations of the aircraft due to flight corrections or rapidly changing wind conditions, e.g. turbulence.

Positions along the line-of-sight $\mathbf{s}$ are denoted by *slant range* r. The slant range of a point $\mathbf{x} = (x, y, z)^{\mathrm{T}}$ in the aircraft-fixed coordinate system is defined as the projection of $\mathbf{x}$ on the line-of-sight $\mathbf{s}$,

$$r(\mathbf{s}, \mathbf{x}) = \mathbf{s} \cdot \mathbf{x}. \tag{2.3}$$

Using the definitions of aspect angle in equations (2.1) and (2.2), the line-of-sight $\mathbf{s}$ can be expressed as a function of aspect angle as

$$\mathbf{s}(\boldsymbol{\theta}) = (-\cos\theta^a \sin\theta^e,\ -\sin\theta^a \sin\theta^e,\ -\cos\theta^e)^{\mathrm{T}}. \tag{2.4}$$

Consequently, the slant range of a point $\mathbf{x}$ as a function of aspect angle $\boldsymbol{\theta}$ is given by

$$r(\boldsymbol{\theta}, \mathbf{x}) = \mathbf{s}(\boldsymbol{\theta}) \cdot \mathbf{x}. \tag{2.5}$$

Suppose that an aircraft, observed initially at aspect angle $\boldsymbol{\theta}$, performs a rotation such that the aspect angle changes by a small amount $\Delta\boldsymbol{\theta}$. Then, the slant range of a point $\mathbf{x}$ can be approximated by a linear function of $\Delta\boldsymbol{\theta}$:

$$r(\boldsymbol{\theta}, \mathbf{x}) \approx r(\boldsymbol{\theta}_0, \mathbf{x}) + r_\perp^a\, \Delta\theta^a + r_\perp^e\, \Delta\theta^e, \tag{2.6}$$

where the *cross ranges* $r_\perp^a$ and $r_\perp^e$ are defined as

$$r_\perp^a = \frac{d}{d\theta^a} r(\boldsymbol{\theta}, \mathbf{x}) = (\mathbf{e}_z \times \mathbf{s}) \cdot \mathbf{x}, \tag{2.7}$$

$$r_\perp^e = \frac{d}{d\theta^e} r(\boldsymbol{\theta}, \mathbf{x}) = \frac{1}{\sin\theta^e} (\mathbf{s} \times (\mathbf{e}_z \times \mathbf{s})) \cdot \mathbf{x}, \tag{2.8}$$

i.e., $r_\perp^a$ is the projection of $\mathbf{x}$ on the direction perpendicular to both the positive z-axis and the line-of-sight, and $r_\perp^e$ is the projection of $\mathbf{x}$ on the direction perpendicular to both the line-of-sight and the plane spanned by the positive z-axis and the line-of-sight.

2.3 Radar Range Profiles

A radar measures the radar reflectivity of an aircraft by transmitting an electro-magnetic pulse and recording the amplitude and phase of the reflected signal (the *radar return*), as a function of time.[2] A *range profile* is defined as the squared amplitude of the radar return. A range profile thus represents the energy of the reflected signal as a function of slant range r.

If the width of the transmitted pulse is much shorter than the size of the aircraft, the different parts of the aircraft contributing to the total radar return are resolved in the resulting range profile, as can be seen in figure 1.1. The width of the transmitted pulse thus determines the resolution of a range profiles – the shorter the pulse, the more detail will be seen in the range profile.

The minimum pulse width a radar can achieve is limited by the power (energy per unit of time) it is capable of transmitting. Unfortunately, most radars are incapable of transmitting the power required to achieve sufficient range resolution.

[2] Throughout this thesis it is assumed that the same radar is used for both transmitting and receiving (monostatic radar).

Pulse compression [24] allows a radar to achieve high range resolution even with relatively long pulse widths. This is achieved by modulating the frequency $f(t)$ of the transmitted pulse during the total pulse width $\tau = t_e - t_s$, where t_s and t_e denote the times of respectively starting and ending the pulse transmission. Radars utilising pulse compression are essentially measuring the total radar return $G(f)$ in the frequency domain. A range profile is then constructed by applying an inverse Fourier transform to $G(f)$. The range resolution of the resulting range profile is determined by the total bandwidth B spanned by $f(t)$.

The *chirp* waveform is a common application of this idea. The chirp waveform is constructed by linearly varying the transmitted frequency over the time period τ,

$$f(t) = f_0 + \frac{t - t_s}{\tau} B, \quad \text{for } t_s < t < t_e. \tag{2.9}$$

Another common waveform [25] is the *stepped frequency* waveform (SFW). Stepped frequency waveforms are produced by linearly sampling the desired bandwidth B at specific frequencies f_l,

$$B = L\Delta f, \tag{2.10a}$$
$$f_l = f_0 + l\Delta f, \tag{2.10b}$$
$$l = 0, \ldots, L. \tag{2.10c}$$

For each frequency f_l, the total radar return $G_l = G(f_l)$ is then measured by transmitting the corresponding pulse at times t_l,

$$t_l = t_s + l\frac{\Delta t}{L}, \tag{2.11}$$

where Δt is the time interval between transmitting two successive pulses.

The corresponding radar return $\mathbf{g}$ in the time domain is then reconstructed by performing an Inverse Discrete Fourier Transform (IDFT) on the $L + 1$ returns G_l,

$$\mathbf{g} = \mathcal{F}^{-1}(\mathbf{G}), \tag{2.12}$$

The elements y_l (usually called *range bins*) of the resulting range profile $\mathbf{y} = (y_0, \ldots, y_L)$ are then given by

$$y_l = |g_l|^2. \tag{2.13}$$

A problem with the use of the SFW is that radial motions of the aircraft cause a distortion of measured range profiles. To minimise these distortions, a *velocity-tolerant* stepped frequency waveform (VTSFW) can be used. In [11] it is shown that by decreasing the time interval between transmitting the frequencies f_l, any constant velocity of the aircraft is automatically compensated for. A VTSFW is produced by transmitting the frequencies f_l at specific times

$$t_l = \Delta t_{\min} \frac{(\gamma l + l)(\gamma L + L - \gamma)}{\gamma l + L}, \tag{2.14}$$

where $\Delta t_{\min} = t_L - t_{L-1}$ is the time delay between the two final pulses, and $\gamma = B/f_0$.

The radar return of a complex target can be described as a coherent sum over the returns of a set of discrete *scatterers* [26]. With this assumption, the total coherent return G for a given

frequency f_l and aspect angle $\boldsymbol{\theta}$ is given by

$$G_l(\boldsymbol{\theta}) = \sum_{m=1}^{M} G_l^m(\boldsymbol{\theta}), \tag{2.15a}$$

$$G_l^m(\boldsymbol{\theta}) = A^m(f_l, \boldsymbol{\theta}) \exp i\phi^m(f_l, \boldsymbol{\theta}), \tag{2.15b}$$

where A^m is the real-valued amplitude and ϕ^m the phase of the return of the mth scatterer. In general, both will be a function of frequency and aspect angle.

The most basic scatterer is the *point scatterer*. A point scatterer is considered to be rigidly attached to the aircraft body at some location $\mathbf{x}$ in the coordinate system shown in figure 2.1.

The radar return of a point scatterer has two defining characteristics: it has a constant amplitude with respect to both frequency and aspect angle, and its phase varies linearly with both frequency f and slant range r. The radar return of a point scatterer located at a position $\mathbf{x}$ in the aircraft-fixed coordinate system is given by

$$G_l(\boldsymbol{\theta}) = A \exp -\frac{4\pi i f_l}{c} r(\boldsymbol{\theta}, \mathbf{x}), \tag{2.16}$$

where $r(\boldsymbol{\theta}, \mathbf{x})$ is defined in equation (2.5).

Many other types of scattering behaviour, such as specular reflections and cavity resonances, are present in actual HRR measurements [27, 28]. However, the relation between position, aspect angle and radar return of these scatter types is very difficult to determine, although efforts have been made to construct parametric models of non-point scattering [29, 30, 31].

It is instructive to explicitly calculate the range profile resulting from a single point scatterer at position $\mathbf{x}$. From equations (2.12) and (2.16) it follows that

$$g_l = A \left(\frac{\sin \pi l'}{\sin \frac{\pi l'}{L+1}} \right) \exp i \left(\phi_l - \frac{4\pi \bar{f}}{c} r(\boldsymbol{\theta}, \mathbf{x}) \right), \tag{2.17}$$

where c is the speed of light, and l', $\bar{f}$ and ϕ_l^0 are defined as

$$l' = l - \frac{2B}{c} r(\boldsymbol{\theta}, \mathbf{x}), \tag{2.18}$$

$$\bar{f} = f_0 + \frac{1}{2} L \Delta f, \tag{2.19}$$

$$\phi_l = \frac{\pi L}{L+1} l. \tag{2.20}$$

The elements y_l of the resulting range profile $\mathbf{y}$ are therefore given by

$$y_l = A^2 \left(\frac{\sin \pi l'}{\sin \frac{\pi l'}{L+1}} \right)^2. \tag{2.21}$$

Equations (2.21) and (2.18) show that a point scatterer causes a local maximum of $\mathbf{y}$ in the range bin y_l that contains the point scatterer. In other words, the slant range of a point scatterer can be estimated from a range profile by locating the local amplitude maxima.

From equations (2.21) and (2.18), two important parameters can be obtained. The *nominal range resolution* Δr is defined as the distance in metres between to successive range bins y_l, and is given by

$$\Delta r = \frac{c}{2B}. \tag{2.22}$$

Due to the periodicity of (2.21), a shift in slant range of an integer multiple of $\frac{c}{2\Delta f}$ will cause the scatterer to appear in the same bin in the range profile. The length

$$R_u = \frac{c}{2\Delta f} \tag{2.23}$$

is therefore called the *unambiguous range interval*.

2.4 Range Profile Variability

This section discusses the main sources of range profile variability: measurement noise, translational range migration, rotational range migration, speckle, and occlusion.

Radar measurements are subject to *measurement noise*, which is caused by both thermal noise in the radar receiver and *clutter* – unwanted radar returns from for example birds or atmospheric effects. The signal-to-noise ratio is mainly determined by the distance between the radar and the aircraft, since the power of the reflected signal measured by the radar is inversely proportional to the fourth power of the distance between the radar and the aircraft.

A range profile suffers from *translational range migration* (TRM) when a change in distance between the radar and the aircraft causes scatterers to move from one range bin to the next. Since all scatterers are translated by the same amount, the relative distance between two scatterers does not change. Therefore, the shape of the profile does not change due to TRM, and so the effect of TRM is a translation of the original range profile. In case of a stepped frequency waveform, this shift is cyclic.

If an aircraft rotates over a significant aspect angle (of the order of a few degrees) such that the outermost scatterers move from one range bin to the other, the range profiles collected during this rotation suffer from *rotational range migration* (RRM) [13].

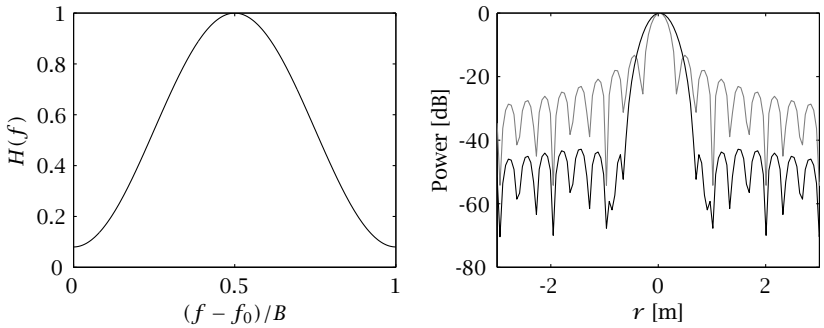
The next source of variability, *speckle*, is also related to aircraft rotations. Speckle occurs if in a single range bin l two or more distinct scatterers are present. Then, only a slight rotation of the aircraft in aspect azimuth or elevation is enough to change the differential path length to the radar over half the wavelength. This causes the coherent sum g_l of the scatter contributions to turn from constructive to destructive interference (or vice versa) within tiny changes of aspect angle; generally between one and two orders of magnitude smaller than the aspect angle changes associated with RRM.

The effect of speckle is that the range profile amplitudes y_l vary rapidly if a sequence of consecutively measured range profiles is considered. In this case, the change in aspect angle is due mainly to small aircraft yaw motions during the recording time. Because the rotations causing speckle are smaller than the accuracy of any aspect angle estimate, speckle is is usually modelled as statistical variance of peak amplitude.

The so-called *Swerling cases* [32] are models of amplitude fluctuations in the case of a very large number of point scatterers being present in a single range bin l. Swerling cases 1 and 2 apply when all scatterers are of comparable strength. Swerling cases 3 and 4 apply when one scatterer is significantly stronger than the other scatterers. Although the Swerling cases are derived from the limit case of an infinite number of scatterers, in practice they work well even if only a few scatterers are present [24].

It has been shown [33] that, for any Swerling case, the resulting distribution $p(y_l)$ of y_l is a chi-square distribution with $2k$ degrees of freedom, given by

$$p(y_l) = \frac{k}{(k-1)!\,\bar{y}_l} \left(\frac{k y_l}{\bar{y}_l}\right)^{k-1} \exp\left(-\frac{k y_l}{\bar{y}_l}\right), \tag{2.24}$$



(a) Frequency domain. (b) Range domain.

Figure 2.2: The Hamming window in both the frequency and the range domain. The gray line in figure (b) shows the IDFT of the default square window.

where $\bar{y}_l$ denotes the average value of y_l. Swerling cases 1 and 2 are described by $k = 1$, while Swerling cases 3 and 4 are described by $k = 2$.

From (2.24) one can calculate the expected variance σ_l^2 on range profile amplitudes y_l. For both $k = 1$ and $k = 2$, this variance varies with y_l as

$$\sigma_l^2 \propto y_l^2. \tag{2.25}$$

Fluctuations of range profile amplitude due to speckle are thus multiplicative in nature – the larger the profile amplitude, the larger the variance. This poses a problem for many statistical classifiers, which are often based on additive noise models. Section 2.5 provides a solution to this problem by applying a non-linear transformation to the range profiles.

The final source of range profile variability discussed here is *occlusion*. Occlusion occurs when a scatterer is positioned such that it is not observable by the radar. In our application, this is always a case of *self-occlusion*, which means a scatterer can only be occluded by another part of the same aircraft. An occluded scatterer does not contribute at all to the measured range profile. Rotations of the aircraft in the order of 10 degrees can cause occluded scatterers to become visible and vice versa.

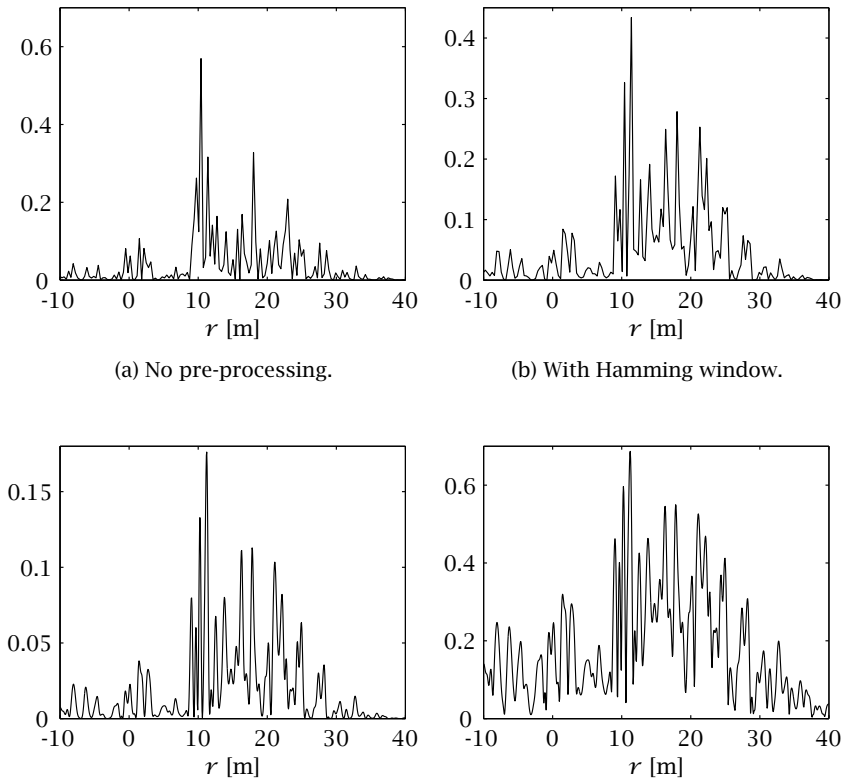
2.5 Range Profile Pre-Processing for Classification

A consequence of using a stepped frequency waveform is that effectively a square window is applied to the 'true' radar return in the frequency domain. This introduces high side-lobes in the range domain. It is therefore common to apply a filter to the complex radar returns before performing the IDFT, which reduces the height of the side-lobes at the cost of a slight loss in resolution.

The choice in filters is rather large (see [34] for an overview), each with their own advantages and disadvantages. The profiles used for the work described in this thesis have all been filtered using the Hamming window [35], shown in figure 2.2. Figure 2.3(b) shows the effect of applying the Hamming window on the resulting range profile. Application of the Hamming window increases the nominal range resolution by roughly a factor of 1.3,

$$\Delta r \approx 1.3 \frac{c}{2B}. \tag{2.26}$$

After applying a window to the $(L+1)$ coherent returns, they can also be zero-padded, thereby



(a) No pre-processing.

(b) With Hamming window.

(c) With Hamming window and zero-padding.

(d) With Hamming window, zero-padding and Box-Cox transform.

Figure 2.3: Examples of the effect of the different pre-processing methods described in section 2.5 on a (measured) range profile. Note that the profiles are normalised to unit vector length, and so the vertical axes are dimensionless.

increasing the dimensionality from $L+1$ to $d \equiv z(L+1)$, where $z > 1$ is the oversampling factor. Zero-padding has the effect of interpolating the resulting range profile, and thereby artificially increasing the number of range bins in the profile. Note that this does not increase the resolution of the profile – it only allows for easier detection of amplitude peaks. The effects of zero-padding on a range profile are shown in figure 2.3(c).

In section 2.4 it was shown (equation (2.25)) that the variance of peak amplitudes due to speckle was multiplicative in nature. However, many classifiers, such as the nearest neighbour classifier, assume an additive noise distribution, such as Gaussian distributed white noise. It can therefore be beneficial to transform the data in such a way that the resulting variance is additive instead of multiplicative.

It has been shown that applying a Box-Cox transformation [36, 37] to range profiles, the multiplicative peak amplitude variation induced by speckle can be transformed to be as additive as possible [11, 38]. A number of equivalent definitions of the Box-Cox transformation are found in literature – in this thesis, the Box-Cox transformation of a range profile $\mathbf{y}$ is defined as

$$y'_l = \frac{(y_l)^\eta}{\sqrt{\sum_d (y_l)^{2\eta}}}, \tag{2.27}$$

where $\mathbf{y}'$ is the transformed profile, and $\eta = [0, 1]$ is a free tuning parameter. The effect of applying a Box-Cox transformation to a range profile is shown in figure 2.3(d). In [11, 38] it has been shown that a value of $\eta \approx 0.2$ is optimal for transforming the multiplicative amplitude fluctuations caused by speckle to normally distributed fluctuations. Note that the transformed profile is also normalised to unit vector length to make the final classification independent on the absolute magnitudes of the range profiles.

2.6 Simulated HRR

Simulated range profiles are produced by supplying radar simulation software with CAD models of aircraft. Simulated range profiles are a cheap and efficient method for obtaining large amounts of profiles, which can be used as a training set for constructing a classifier (as discussed in section 2.1).

Aircraft CAD models represent the geometry of aircraft as a collection of discrete elements. These elements can be parametric functions (such as for example NURBS [39]), or simple primitives (such as the flat patches of the model shown in figure 1.2).

Simulated range profiles are immune to some of the sources of range profile variability discussed in section 2.4. Simulated range profiles do not suffer from measurement noise (apart from negligible errors due to the limited numerical accuracy of computer computations). Also, simulated range profiles do not suffer from translational range migration (provided the origin of rotations of the aircraft model is fixed).

Other differences between measured and range profiles result from the limited accuracy of HRR simulation. First of all, simulation software can only approximate the process of radar scattering. Also, CAD models are only approximations of the geometry of aircraft.

A large variety of radar simulation software packages exists. A frequently used package is XPATCH [40]. For the work presented in this thesis, the RAPPORT[3] software package [41] was used exclusively. This section will therefore restrict the discussion of radar simulation to the approach taken in RAPPORT.

RAPPORT is a radar simulation package which is developed at TNO Physics and Electronics Laboratory. It calculates the total coherent radar return (amplitude and phase) of an aircraft model (for a given frequency and aspect angle) using a combination of ray tracing and *physical optics*. RAPPORT uses a ray tracing algorithm to calculate the trajectory of the radar waves as they hit the aircraft. From these paths (which may contain multiple bounces), it determines which of the flat patches in the aircraft model are 'visible' to the radar, i.e. contribute to the total radar return.

RAPPORT then uses a high frequency approximation to radar scattering called *physical optics* to calculate the radar return of each visible facet. The total radar return is the coherent sum of the contributions of each facet.

2.7 Available Data

The data set used throughout this thesis contains both simulated and measured HRR data collected from five civil aircraft, seen at approximately broadside aspect angles: the Boeings 737-500 and the 747-400, the Fokker 100, the Airbus A310 and the McDonnell-Douglas of the 80–88 series.

The measured range profiles were collected during the ORFEO measurement campaign [42], using the FELSTAR[4] radar at TNO Physics and Electronics Laboratory. The range profiles were

[3] **R**adar signature **A**nalysis and **P**rediction by **P**hysical **O**ptics and **R**ay **T**racing
[4] **FEL S**-band **T**racking and **A**cquisition **R**adar.

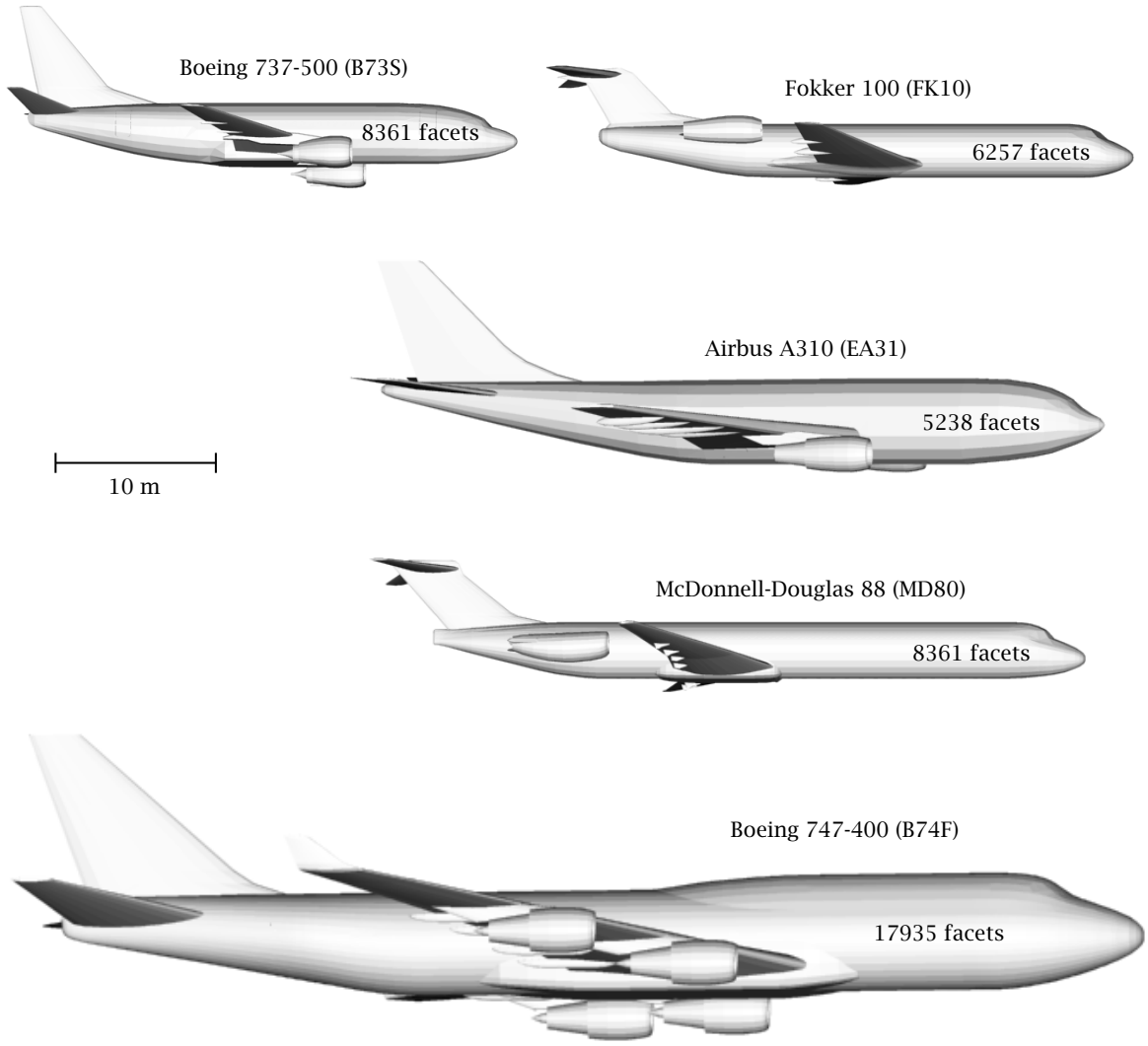


Figure 2.4: The five models available for RCS-predictions. Above each model the name of the aircraft type is shown. Between brackets, the code-names used throughout this thesis are given. The number of flat facets ranges from 5,238 for the Airbus to 17,935 for the Boeing 747.

Carrier frequency	f_0	=	3.0738 GHz
Bandwidth	B	=	452.2 MHz
Number of pulses	$L+1$	=	324
Frequency step	Δf	=	1.4 MHz
Nominal range resolution	ΔR	=	0.33 m
Unambiguous range	R_u	=	107.5 m
Minimum delay between pulses	Δt_{min}	=	420 μs

Table 2.1: ORFEO waveform parameters

measured using the velocity tolerant stepped-frequency waveform described in section 2.3. Table 2.1 shows the parameters of the waveform used.

A typical ORFEO measurement consists of multiple profiles of an aircraft in flight (i.e. a leg) as the full waveform is repeatedly transmitted while the radar is tracking the aircraft. For each leg, aspect angle estimates were obtained from the tracking data. The aspect angle estimates contain a systematic error of about 5 degrees, and a random error of about one degree. More detail on the ORFEO measurement campaign is provided in [11, 42].

For each aircraft in the database a CAD model was used to obtain simulated profiles, using the RAPPORT software package. Each CAD model consists of a collection of flat patches (or *facets*), whose union approximates the shape of the aircraft. Figure 2.4 shows a rendered representation of these models. The number of flat facets ranges from 5,238 for the Airbus to 17,935 for the Boeing 747. This demonstrates that the number of facets for the Airbus is relatively low. From figure 2.4 the rougher surface of the Airbus can actually be observed. Clearly, the predictions for the Airbus will therefore be less accurate than for the other aircraft types.

2.8 Conclusions

This section provided an overview of the basic physical and statistical properties of radar range profiles. Radar range profiles, denoted by $\mathbf{y}$. are measurements of reflected radar energy as a function of slant range r.

The measured range profiles used throughout this thesis are obtained using a Velocity Tolerant Stepped Frequency Waveform. The benefits of this waveform are twofold. First, by transmitting the total required bandwidth over a relatively large period Δt, the hardware requirements on the radar equipment are modest. Second, by a specific choice for the times at which the specific frequencies in the waveform are transmitted, the resulting measurements are not influenced by radial motions of the aircraft. All measured range profiles used in this thesis were obtained during the ORFEO measurement campaign.

The main sources of range profile variability are measurement noise, translational range migration, speckle, rotational range migration and occlusion. Translation range migration occurs when the distance between the radar and the aircraft changes due to aircraft motions, and causes cyclic shifts of range profiles. Speckle, rotational range migration and occlusion are all effects caused by changes in aspect angle $\boldsymbol{\theta}$. Speckle causes rapid amplitude fluctuations when more than one scatterer are present in a single range bin. RRM causes scatterers to move from one range bin to the next. Occlusion causes scatterers to be 'invisible' in range profiles, and occurs when a scatterers are hidden from the radar by other parts of the aircraft.

The usefulness of radar range profiles for classification can be enhanced using signal processing. By windowing the radar returns in the frequency domain, high side lobes in the range domain are suppressed. Zero-padding the returns allows for easy interpolation of the range profile, which allows for more accurate detection of amplitude peaks. Finally, the Box-Cox

transformation transforms the multiplicative variability caused by speckle into additive variability.

Simulated range profiles are produced using radar simulation software with CAD models of aircraft. They are free of the influence of measurement noise and TRM. Due to limitations of both the simulation software and the CAD models, they are necessarily only an approximations to measured range profiles. All simulated range profiles used in this thesis were obtained using the RAPPORT software package.

3 Translation Invariant Classification of Radar Range Profiles[1]

3.1 Introduction

Measured radar range profiles are subject to the effects of translational range migration (TRM): changes in the distance between the radar and the aircraft causes cyclic shifts of the range profiles. Since the distance between radar and aircraft cannot be measured with sufficient accuracy to correct for this effect, measured range profiles are subject to arbitrary shifts. This is a problem for classifiers, since two or more range profiles can only be meaningfully compared after they are *aligned*, either with respect to each other, or with respect to some external reference frame. Another solution is to classify range profiles on the basis of translation invariant features extracted from the range profiles

A general classification framework consists of several discrete stages: *data acquisition* (measuring radar returns, computing profiles within a simulated environment), *pre-processing* (scaling, noise removal), *feature extraction* (including dimension reduction) and finally the actual *classifier*. Translation invariance can, in principle, be achieved at any of these stages.

During the pre-processing stage of a classification method, data is prepared for later feature extraction. Pre-processing typically includes transformations such as scaling and noise removal. Translation invariance in this phase is achieved by *registering* the profiles: each profile is translated such that some optimisation criterion is maximised.

In the literature both relative and absolute alignment procedures are found in this stage. It is important to note that this problem is often ignored in literature. Also, many papers report on experiments using simulated profiles only – in this case the exact distance between radar and aircraft always constant, and so TRM is not a source of error.

Relative alignment is commonly achieved by aligning two profiles such that their correlation is maximised. The third profiles is then aligned with respect to the second, the fourth with respect to the third, and so forth. One drawback of these and similar approaches, extensively discussed in [43], is that if one profile is misaligned, this error propagates and disturbs the alignment of subsequent profiles in the data set. In [43] a method is proposed to deal with these misaligned profiles.

A method for obtaining absolute alignment during the pre-processing stage is given in [22]. This method, which the authors call *auto-aligning*, registers profiles by translating them such that the entropy of an energy vector E, whose elements are the inner products between the translated profile and K Gaussian windows with increasing standard deviation σ_k, is minimal.

The second opportunity for obtaining translation invariant classification is during the feature extraction stage. A common approach [12] is to use the magnitude of the Fourier transform of the range profiles as feature vectors, discarding all phase information. In [22] translation invariant features are obtained by first calculating the bi-spectrum (defined as the 2-D Fourier transform of the third-order autocorrelation function) for each profile, which is then integrated and inverse Fourier-transformed to obtain features (called *accumulated bi-spectral features*).

Using translation invariant features can significantly decrease the computational cost of classification. The main disadvantage of these and similar approaches is that extracting translation

[1] A paper containing parts of this chapter has been accepted for publication in *IEE Proceedings – Radar, Sonar and Navigation*.

invariant features involves discarding potentially discriminating information, which can result in a decrease of classification performance. Exactly which features should be considered, and whether this approach in general provides acceptable performance, is still an open question.

Obtaining translation invariant classification in the actual classifier is probably the most common approach to solving the problem of translation invariant classification reported in recent literature. In [11], several similarity-based classifiers using the Sliding Euclidean Distance (SED) are discussed. The SED is a metric defined as the minimum Euclidean distance between two profiles $\mathbf{y}_1, \mathbf{y}_2$ over all possible cyclic translations:

$$\mathrm{SED}(\mathbf{y}_1, \mathbf{y}_2) = \arg\min_l |(S^l \mathbf{y}_1) - \mathbf{y}_2|^2, \tag{3.1}$$

for $l = 1, \ldots, L$, where L is the total number of range bins in each profile. The operator S^l shifts its argument vector l elements cyclically to the left (and so $S^L \mathbf{y} \equiv \mathbf{y}$). Since

$$|(S^l \mathbf{y}_1) - \mathbf{y}_2|^2 = |\mathbf{y}_1|^2 + |\mathbf{y}_2|^2 - 2\,(S^l \mathbf{y}_1) \cdot \mathbf{y}_2, \tag{3.2}$$

the SED is equivalent to the normal Euclidean distance after first aligning $\mathbf{y}_1$ and $\mathbf{y}_2$ such that their cross-correlation

$$\boldsymbol{\lambda}(l) = (S^l \mathbf{y}_1) \cdot \mathbf{y}_2 \tag{3.3}$$

is maximal. The Fast Fourier Transform allows for an efficient implementation of this procedure. Let $\mathbf{Y}$ denote the FFT $\mathcal{F}(\mathbf{y})$ of $\mathbf{y}$. The *discrete correlation theorem* [44] states that

$$\boldsymbol{\lambda} = \mathcal{F}^{-1}\left[\mathbf{Y}_1 \mathbf{Y}_2^*\right], \tag{3.4}$$

where $\mathbf{Y}_2^*$ is the complex conjugate of $\mathbf{Y}_2$.

In [14, 21] similar approaches using cross-correlation are described. Some disadvantages of these approaches are summarised in [43]. A general drawback of correlation-based translation-invariant classifiers (including the SED) is that they are computationally expensive, since a correlation has to be computed between each range profile to be classified and every range profile in the training set.

In [18] translation invariance is achieved in a model-based classifier by modelling the statistical variations of profile centroids with respect to a set of reference profiles. During actual classification, the likelihood that an unknown test profile $\mathbf{y}$ belongs to a given class C is integrated over all possible centroid positions, weighted by the centroid distribution.

A general disadvantage of obtaining translation invariance in the actual classification stage is that meaningful feature extraction becomes very difficult. Many often-used statistical feature extraction methods (like Principal Component Analysis) depend heavily on statistical properties of the data. Incorrect alignment will affect these properties, decreasing the accuracy of the feature extraction, which in turn decreases classification performance.

This chapter introduces a method for obtaining translation invariant range profile classification by absolute alignment in the pre-processing stage. Section 3.2 shows how the problem of determining an absolute alignment of radar range profiles can be described as a problem of phase estimation. Section 3.2.2 describes an alignment method based on this description, called the Smoothed Zero Phase Representation. Experiments to quantify the effect of different alignment methods on classification performance are described in section 3.3, the results of which are presented and discussed in section 3.4. Finally, section 3.5 draws conclusions from the experimental results.

3.2 Zero Phase Representation

The effects of Translational Range Migration can be best described in terms of the phases of the Fourier-transformed profiles. A well-known symmetry of the Fourier transform is that for

any function $f(x)$ with Fourier transform $\mathcal{F}(f)$, the Fourier transform of $f(x-s)$ is given by

$$\mathcal{F}(f(x-s)) = e^{i\omega s}\ \mathcal{F}(f(x)), \tag{3.5}$$

i.e. a translation in the range domain is equivalent to a phase shift of $F(f(x))$ in the frequency domain. In the discrete case, where the function f is represented as a set of samples $f_n = f(x_n)$ for $n = 1, \ldots, d$, a similar property holds for cyclically shifted versions of f_n:

$$\mathcal{F}(f_{(n+k) \bmod d}) = \mathcal{F}(f_n) e^{2\pi i n k/d}. \tag{3.6}$$

So, for a discrete shift k, the phase ϕ of the first AC component will be shifted by $2\pi k/d$, and higher order phases by integer multiples of $2\pi k/d$.

Suppose a leg consisting of N profiles $\mathbf{y}_i$, $i = 0, \ldots, N-1$ is measured from an aircraft in flight at discrete times t_i. Let $\phi(t_i) \equiv \phi_i$ denote the corresponding phases of the first AC component of the FFT's of $\mathbf{y}_i$.

Variation in ϕ_i is caused by both rotational and translational effects. Suppose for the moment that the aircraft remains at a fixed distance R w.r.t. the radar while rotating. This means ϕ_i is solely a function of the aspect angles θ^a and θ^e:

$$\phi_i = \phi(\theta^a(t_i), \theta^e(t_i)) \equiv \phi_i^a. \tag{3.7}$$

Since the aircraft remains at a fixed position, all variation in ϕ is the result of speckle and rotational range migration.

In a more realistic setting, the distance $r(t_i) \equiv r_i$ between radar and aircraft also varies with time. This is a linear effect on ϕ_i,

$$\begin{aligned} \phi_i &= \phi_i^a + \frac{2\pi}{R_u}(r_i - R) \\ &= \phi_i^a + \beta(r_i - R), \end{aligned} \tag{3.8}$$

where $\beta = 2\pi/R_u$ and R_u is the unambiguous range interval (defined in section 2.3).

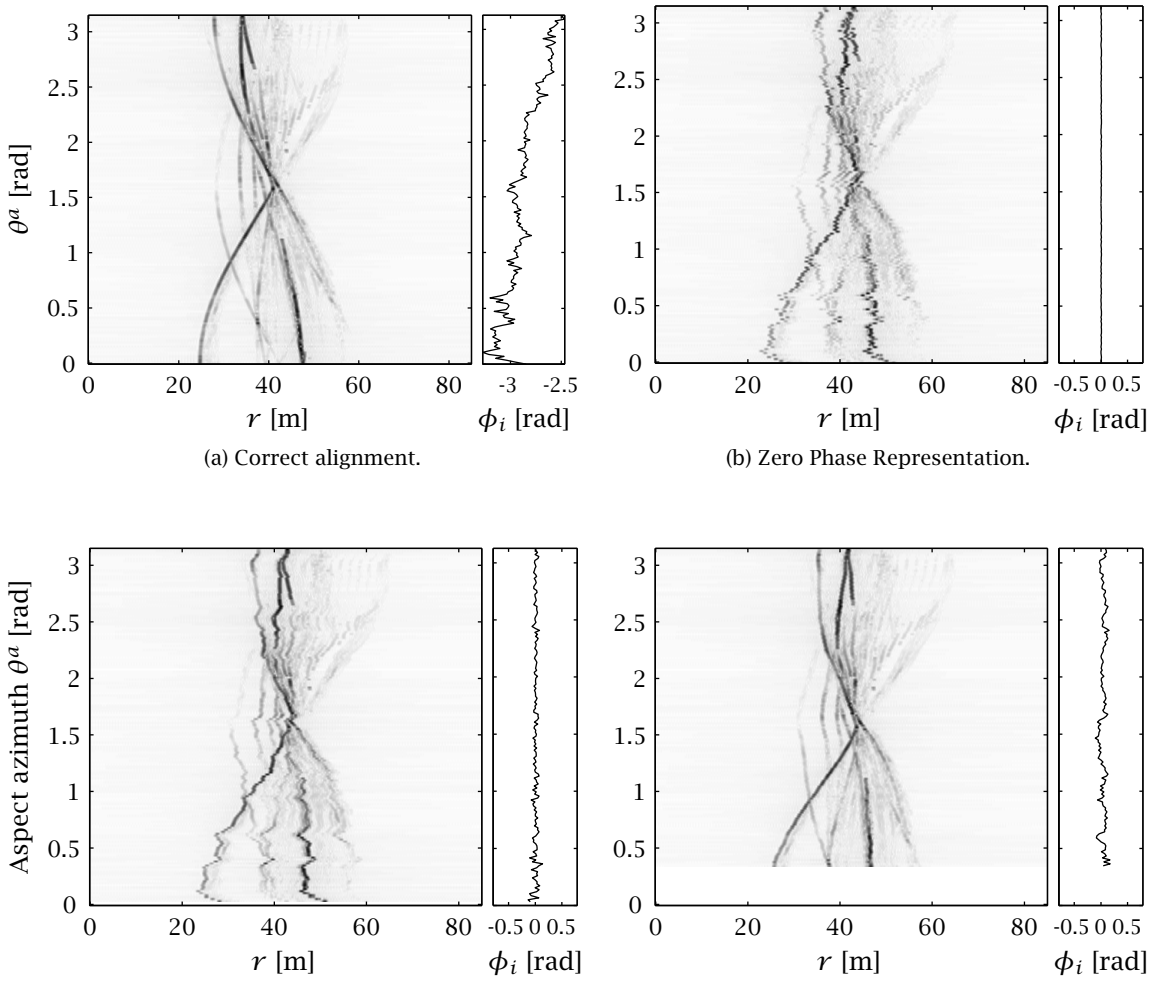
The last term in (3.8) is the source of TRM. If the r_i were known exactly, reversing the effects of TRM would be trivial: the phase sequence ϕ_i could be adjusted by subtracting $\beta(r_i - R)$ at each time step, and the higher phases by integer multiples of $\beta(r_i - \mathbf{R})$. However, r_i is not known, and so the last term in (3.8) has to be estimated through other means.

3.2.1 Pure Zero Phase Representation

A simple method for finding a translation invariant representation (which has been previously used for alignment of panoramic images [45]), is setting each ϕ_i to zero and adjusting the higher order phases accordingly. This eliminates the range term in (3.8), but at the same time introduces an alignment error due to discarding the ϕ_i^a.

As an example, in figure 3.1(a) a leg of simulated range profiles of a Fokker 100 aircraft viewed over a 180 degree turn is shown. These profiles are perfectly aligned as the aircraft rotated around its center without changing its distance to the radar. For comparison the same leg in the Pure Zero Phase Representation (ZPR) is shown in figure 3.1(b).

From figure 3.1 it is clear that the ZPR results in a very rough alignment. The clear tracks visible in figure 3.1(a) are still recognisable but less smooth in figure 3.1(b).



(a) Correct alignment.

(b) Zero Phase Representation.

(c) Smoothed ZPR, $w = 20$.

(d) Smoothed ZPR, $w = 200$.

Figure 3.1: Outputs of the several alignment procedures for one leg of simulated profiles. The profiles were obtained at aspect elevation $\theta^e = \pi$, and aspect azimuth θ^a ranging from 0 to π. On the right, the phases of the resulting profiles are plotted.

3.2.2 Smoothed Zero Phase Representation

This section proposes a more sophisticated method for dealing with the range terms in (3.8): the Smoothed Zero Phase Representation (SZPR). It is based on the assumption that the relative location of the majority of prominent scatterers within the profile will remain stable over small changes in aspect angle. Therefore, two successive profiles in a leg can be accurately aligned *relative to each other* by maximum correlation (as in equation (3.3)). Relative alignment through maximal correlation is a popular approach in the literature for obtaining translation invariant classification within similarity-based classifiers.

The goal of the SZPR is to combine the advantages of relative alignment through maximal correlation with the advantages of an absolute alignment using the pure ZPR. Consider again equation (3.8). The phases ϕ_i and ϕ_{i-1} measured at times t_i and t_{i-1} are related through

$$\begin{aligned} \phi_i - \phi_{i-1} &= \phi_i^a - \phi_{i-1}^a + \beta(r_i - r_{i-1}) \\ &= \phi_i^a - \phi_{i-1}^a + \beta\Delta r_i, \end{aligned} \tag{3.9}$$

for $i = 0, \ldots, N-1$. Now, for each profile $\mathbf{y}_i$, $i > 0$, the optimal shift with respect to the previous profile $\mathbf{y}_{i-1}$ can be estimated using using maximal correlation. This shift corresponds to a phase shift $\delta\phi_i^c$. The assumption that this phase shift correctly aligns profile $\mathbf{y}_i$ with respect to $\mathbf{y}_{i-1}$ is equivalent to the assumption $\delta\phi_i^c = \beta\Delta r_i$. Since β is known, this means that for each measurement (except the first) the relative translation Δr_i of the aircraft between two successive measurements can be estimated.

Using this estimate, a corrected phase sequence ϕ_i^c can be constructed from the original phases ϕ_i as

$$\begin{aligned} \phi_i^c &= \phi_i^a + \beta(r_i - R - \sum_{j=1}^{i} \Delta r_j) \\ &= \phi_i^a + \beta(r_0 - R) \end{aligned} \tag{3.10}$$

for $i = 1, \ldots, N-1$. In other words, all profiles are now registered at $r = r_0$, the distance between radar and aircraft at the moment the first measurement was made.

To obtain absolute alignment, all profiles should be registered at the reference distance R. This can only be achieved if r_0 is known. However, r_0 is not known, but it can be estimated using the following heuristic approach. At each time-step, $\beta(r_0 - R)$ is estimated *locally* by a moving average over ϕ_i^c. This estimate is then subtracted from ϕ_i^c to obtain a final phase estimate ϕ_i^f:

$$\phi_i^f = \phi_i^c - \frac{1}{w+1} \sum_{k=-w}^{w} \phi_{i+k}^c. \tag{3.11}$$

By varying the window size w of the moving average filter, this alignment procedure can be 'interpolated' between the absolute alignment of the pure ZPR ($w = 1$) and relative alignment through pure correlation ($w = N$).

In theory the optimal window size w depends on the differences in aspect angle between successive profiles, which in turn depends on the pulse repetition frequency (PRF) and the target rotation rate. For small differences in successive aspect angles (high PRF and/or low target rotation rate), alignment by maximum correlation works well, and so a large value of w can be used. When the PRF is small, or the target rotates quickly, the quality of alignment by maximum correlation degrades, and and so a smaller value of w should be used.

However, in practical situations, i.e. when looking at targets of opportunity, the target rotation rate can not be measured accurately enough to allow for dependable estimates of w. This

chapter therefore tries to find a globally acceptable value for w by experimentation.

Figures 3.1(c) and 3.1(d) show an example of the effect of the SZPR on a stack of profiles. Since filtering with a windows size w makes aligning the first $w-1$ profiles impossible, these are left blank in the plots.

3.2.3 Implementation and Computational Cost

This section compares the computational cost of nearest neighbour classification using the SED with classification using the SZPR and a normal, i.e. non-sliding, Euclidean distance. The example task is the classification of a leg consisting of M profiles $\mathbf{y}_m$, given a training set containing N pre-aligned (using the SZPR) training profiles $\mathbf{y}_m$.

Nearest neighbour classification using the SED requires following computations for each of the MN possible pairs of test profiles $\mathbf{y}_m$ and training profiles $\mathbf{y}_n$:

1. Calculate the FFT's $\mathbf{Y}_m$ and $\mathbf{Y}_n$ of $\mathbf{y}_m$ and $\mathbf{y_n}$.
2. Calculate the correlation function $\boldsymbol{\lambda} = \mathcal{F}^{-1}(Y_n Y_m^*)$ and find the maximum of $\boldsymbol{\lambda}$. This determines the optimal alignment from maximum correlation.
3. Collect the phase ϕ_m of the first AC component of the $\mathbf{Y}_m$'s, and adjust it such that it reflects the optimal alignment.
4. Use the adjusted phase to align $\mathbf{Y}_m$ with $\mathbf{Y}_n$.
5. Calculate the Euclidean distance between the two profiles (which can be done directly in the Fourier domain.)

Nearest neighbour classification in the SZPR, using a normal Euclidean distance measure, requires the following computations to be performed:

1. Calculate the FFT's $\mathbf{Y}_m$ and $\mathbf{Y}_n$ of $\mathbf{y}_m$ and $\mathbf{x_n}$.
2. Calculate the $(M-1)$ correlation function $\boldsymbol{\lambda} = \mathcal{F}^{-1}(Y_m Y_{m-1}^*)$ and find the maximum of $\boldsymbol{\lambda}$. This determines the optimal alignment from maximum correlation.
3. Collect the phases ϕ_m of the first AC component of the $\mathbf{Y}_m$'s, and adjust them such that the leg is now aligned using maximum correlation.
4. Filter the resulting phase sequence using a moving average filter with window size w.
5. Use the the filtered phase sequence to align $\mathbf{Y}_m$ in the SZPR.
6. Calculate the Euclidean distance between each of the MN possible pairs of test and train profiles (which can be done directly in the Fourier domain).

Table 3.1 summarises the computational costs, showing the number of times different computations have to be performed. From table 3.1 it is clear that classification using the SZPR is significantly faster than using the SED. The computations directly associated with aligning the profiles have to be computed N times less when using the SZPR than when using the SED. The filtering of the phase sequence adds some computational cost to the SZPR, but the filtering can be implemented efficiently using the Fourier transform, similar to the correlation function [44], and does not contribute significantly to the total number of computations performed.

Task	SED	SZPR
FFT	MN	MN
Calculating $\boldsymbol{\lambda}$	MN	$M-1$
Phase adjustments	MN	$M-1$
Filtering phase sequence	-	1
Shifting profiles	MN	M
Euclidean distance	MN	MN

Table 3.1: Computational cost of classification using the SED and the SZPR. The table elements on the right denote the number of times the computations shown on the left have to be performed to classify a leg containing M profiles on the basis of a training set containing N profiles.

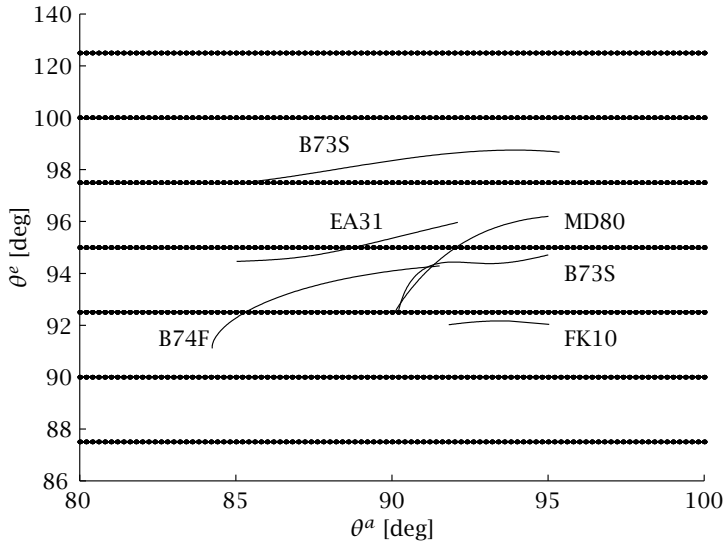


Figure 3.2: The aspect angles of the measured and predicted range profiles. The solid lines represent the measurements. Each curve is labelled by the call sign of the aircraft (defined in section 2.7). Each of the dots represent the aspect angle of five predicted range profiles, one for each aircraft type.

3.3 Experiments

3.3.1 Available Data

For the classification experiments, two sets of range profiles are used in this chapter. The first set contains *measured* range profiles, the second set contains *simulated* range profiles.

The measured data set contains a total of 836 profiles in 6 legs. The measurements were performed during the ORFEO measurement campaign, discussed in section 2.7. The estimated aspect angles are plotted in figure 3.2 as well.

For each aircraft simulated profiles were calculated at 505 grid-points in aspect azimuth and elevation: one profile every 2.5 degrees in aspect elevation and every 0.2 degrees in aspect azimuth. See figure 3.2. All simulations were done using the RAPPORT software package, discussed in section 2.7.

All range profiles in both data sets were Hamming weighted and zero-padded to 512 range bins. Each range profile $\mathbf{y}$ was consequently transformed using the Box-Cox transformation defined in equation (2.27), with $\eta = 0.2$.

3.3.2 Construction of Simulated Legs

In the case of measured profiles, the concept of a leg is clear: profiles are collected from each aircraft a number of times in rapid succession. Each set of measurements is therefore a leg: a sequence of profiles ordered in time.

However, simulated profiles are only labelled by aspect angle. Since the SZPR relies on ordered sequences of profiles, artificial legs have to be constructed from the set of simulated profiles. For each aircraft in the simulated data set, first all profiles obtained at identical aspect elevation are collected. These collections are then ordered according to aspect azimuth. Therefore, one leg of simulated data consists of 105 profiles at a constant aspect elevation, ordered by aspect azimuth which ranges from 80 to 100 degrees with a step size of 0.2.

3.3.3 Range Profile Classification

For the purpose of comparing the quality of different alignment methods the choice of classifier is arbitrary. This chapter therefore solely uses the 1-nearest neighbour classifier, since it is is straightforward to implement, and appropriate for all considered alignment methods. All classification experiments reported in this chapter are therefore based on this algorithm.

To examine the effects of the different alignment methods on classifier performance (expressed as the percentage of correct classifications), two sets of experiments were performed.

For the *first* set of experiments the set of simulated profiles was split into a training and a test set. The training set was produced by taking the odd numbered profiles from each leg. The test set consisted of all remaining, even-numbered profiles. In reality the training and test set are not so highly correlated, and consequently the classification performance will in general be lower than the performance reported in this chapter. However, the goal of this chapter is to investigate the relative merits of the various alignment techniques, and therefore absolute classifier performance is not relevant.

The effect of different alignment procedures on the the classification performance was measured using the classification performance on correctly aligned sets as a comparison baseline.

For the *second* set of experiments, the training set consisted of all available simulated profiles. The test set consisted of the legs of *measured* profiles. Again, the effect of different alignment procedures was measured in terms of classification performance.

3.4 Results

3.4.1 Classification Performance

The classification (using a 1-nearest neighbour classifier) of the simulated test data using simulated training data was optimal, i.e. all 1751 test profiles were classified correctly. A perfect classification was also obtained by using the SED instead of the normal Euclidean distance.

Using the pure Zero Phase Representation for absolute alignment results in the confusion matrix shown in table 3.2. In this case, 13 profiles were classified incorrectly. Classification performance in the SZPR was also measured for a window size ranging from 1 to 20. Optimal performance was obtained at a window size of 8. In this case, only one profile was classified incorrectly. These experiments show that in theory the SZPR is a viable method for absolutely aligning range profiles.

In practice, however, one wants to classify measured profiles from a training set of simulated profiles. To mimic this situation (and simultaneously research the effect of noisy data on alignment), artificially generated measurement noise was added to the simulated test set. Gaussian noise with variance σ^2 and zero mean was superimposed on the radar returns in the frequency

Class	B73S	B74F	EA31	FK10	MD80
B73S	**0.99**	0.00	0.00	0.00	0.01
B74F	0.00	**0.99**	0.00	0.00	0.00
EA31	0.00	0.00	**1.00**	0.00	0.00
FK10	0.00	0.00	0.00	**0.99**	0.01
MD80	0.00	0.00	0.00	0.01	**0.99**

Table 3.2: Confusion matrix using a 1-NN classifier on profiles in the Zero Phase Representation. The table shows the fraction of range profiles assigned to the classes shown in the column labels. The row labels denote the true class label. The overall classification error is 1 %.

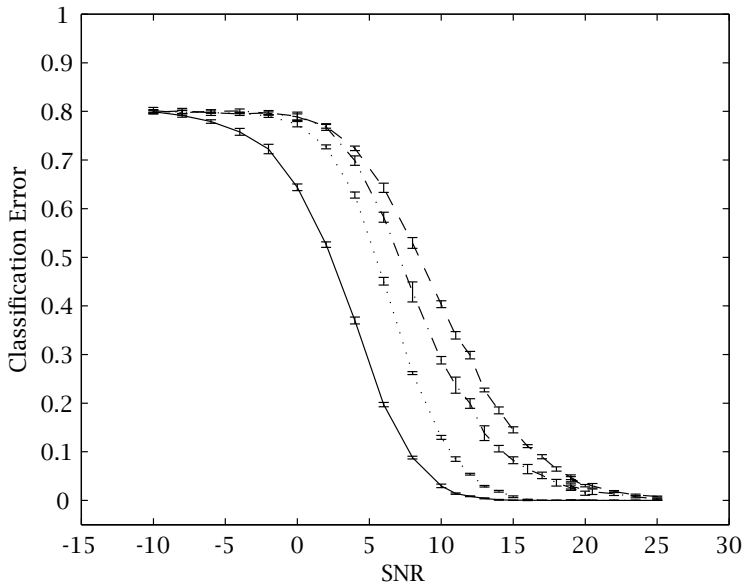


Figure 3.3: Classification error for correctly aligned profiles (solid), using the SED (dotted), the ZPR (dashed) and profiles in the SZPR using a window size $w = 8$ (dash-dotted) as a function of the SNR.

domain of the simulated test set. Then, overall classification performance was measured for correctly aligned profiles, profiles aligned using the SED, and profiles in the ZPR and SZPR as a function of σ. For the SZPR a window size of $w = 8$ was used. As a reference, the results for optimally aligned profiles are provided as well. All results are shown in figure 3.3.

This process was repeated 5 times, each time drawing independent samples from the Gaussian noise distribution. The classification results were averaged over these 5 experiments. The error bar in 3.3 shows the standard deviation over these 5 measurements.

As could be expected, using correctly aligned profiles is optimal in terms of classification performance. Classification using the SED outperforms the SZPR, and is more robust with respect to noise. Recall, however, that classification in the ZPR/SZPR is roughly 500 times faster compared to the SED. Also, observe that for low signal-to-noise ratios, the lines converge to an error of 0.8 since the classification becomes effectively random over five classes.

Finally, classification experiments using the measured profiles for testing were performed. The training set consisted of *all* simulated profiles. Again, the SZPR was tested for window sizes ranging from 1 (which is equivalent to the pure ZPR) to 20. For this set, again a window size $w = 8$ produced the best classification in the SZPR. In table 3.3, confusion matrices for the SED, the ZPR, and the SZPR with $w = 8$ are shown. As in the case where simulated profiles

Class	B73S	B74F	EA31	FK10	MD80
B73S	**0.96**	0.00	0.00	0.04	0.00
B74F	0.00	**0.98**	0.00	0.02	0.00
EA31	0.32	0.00	**0.47**	0.04	0.17
FK10	0.03	0.00	0.00	**0.82**	0.15
MD80	0.09	0.00	0.00	0.89	**0.02**

(a) Confusion matrix for classification using the SED. Overall classification error is 30.4 %.

Class	B73S	B74F	EA31	FK10	MD80
B73S	**0.82**	0.00	0.00	0.16	0.01
B74F	0.00	**0.96**	0.00	0.03	0.01
EA31	0.36	0.00	**0.44**	0.03	0.18
FK10	0.06	0.02	0.00	**0.53**	0.39
MD80	0.16	0.00	0.00	0.83	**0.01**

(b) Confusion matrix for classification in the ZPR. Overall classification error is 38.5 %.

Class	B73S	B74F	EA31	FK10	MD80
B73S	**0.79**	0.00	0.00	0.19	0.02
B74F	0.00	**0.96**	0.00	0.01	0.03
EA31	0.27	0.01	**0.47**	0.02	0.23
FK10	0.02	0.00	0.00	**0.87**	0.11
MD80	0.12	0.00	0.00	0.83	**0.05**

(c) Confusion matrix for classification in the SZPR. The case reported is for a window size of 8. Overall classification error is 35.5 %.

Table 3.3: Confusion matrices for the various alignment methods. The tables show the fraction of range profiles assigned to the classes shown in the column labels. The row labels denote the true class label. The training set consists of synthetic profiles, while the test set consists of measured profiles.

were used for both training and testing, classification using the SED outperforms both the ZPR and the SZPR, although the differences in performance between the three methods are smaller. The significantly overall larger classification error of all methods compared to the two previous experiments are due to the fact that in this experiment, measured range profiles are classified using a simulated training set.

3.5 Conclusions

This chapter described two new methods for aligning range profiles during the pre-processing stage of a classification framework. The first method, called the Zero Phase Representation, obtains absolute range profile alignment by shifting range profiles such that the phase of the first AC component of their Fourier transform is zero. The second method, called the Smoothed Zero Phase Representation, effectively interpolates between the absolute alignment obtained using the pure Zero Phase Representation and the relative alignment obtained by correlating successive range profiles in a leg.

The main benefits of both methods are twofold. First, the range profiles are aligned in a pre-processing stage, resulting in more possibilities for consequent statistical feature extraction. Second, when compared to the common approach to translation invariant classification using the Sliding Euclidean Distance, range profiles aligned using the ZPR or SZPR can be classified much faster, since in this case a nearest neighbour classifier using the standard Euclidean distance measure can be used. Classification of a leg containing M range profiles requires roughly M times less alignment-related computations than classification using the SED.

Using the ZPR or SZPR does result in a decrease in classifier performance compared to classification using the SED. The difference in performance has been shown to be dependent on the signal-to-noise ratio of the range profiles. For both very high signal-to-noise ratios (over 20 dB) and very low signal-to-noise ratios (below 5 dB), the difference in performance is negligible. For intermediate signal-to-noise ratios, classification using the SED outperforms both the SZPR and the ZPR, with the SZPR outperforming the ZPR.

However, by combining range profile alignment using the SZPR with subsequent statistical feature extraction could result in better classification performance. Whether such an approach outperforms classification using the SED would be an interesting topic for future research.

4 Improving Aircraft Rotation Estimates

4.1 Introduction

The aspect angle at which an aircraft is observed heavily influences range profile measurements. The effects of speckle, rotational range migration and occlusion (as was discussed in chapter 2) are the cause of significant range profile variability over both small (less than one degree) and large (tens of degrees) changes in aspect angle. Therefore, knowledge about the aspect angle at which a range profile is obtained can significantly increase classification accuracy. Furthermore, information on aspect angles also benefit many HRR signal processing techniques.

Aspect angles of aircraft can either be measured using onboard equipment, or they can be estimated using tracking data [23]. Onboard equipment is only available in controlled experiments, and so for the purpose of non-cooperative aircraft classification estimating aspect angles from tracking data is the only source of aspect angle information.

Some research reported in the literature treat aspect angle estimation as an integral part of the classification task. In these approaches, the task of the classifier is to provide a joint estimate of class label and aspect angle [46, 47, 48]. Such approaches are useful if no a priori information on aspect angle is available. This chapter assumes that aspect angle estimates from tracking data are available, as is the case for the ORFEO data set described in chapter 2.

Aspect angle estimates from tracking data suffer from both systematic and random errors (see chapter 2). The goal of this chapter is to improve the accuracy of aspect angle estimates by applying a filter to a combination of the original estimates and phase information extracted from radar measurements.

A simple point scatter model of radar scattering provides the basis for the filter. Section 4.2 shows that a simple linear relationship exists between the phase of a point scatterer and the aircraft rotation. This linear relationship is then used in section 4.3 to construct a filter for aircraft rotations.

The derivation of the filter assumes that the phase of a number of point scatterers is extracted from a leg of range profiles. Section 4.4 develops the feature extraction algorithm necessary to extract these phase from the measurements.

Section 4.5 describes the experiments performed to evaluate the improvement of aircraft rotation estimates. Unfortunately, since no onboard equipment was present on the aircraft measured in the ORFEO campaign, the set of available measured range profiles available for this thesis can not be used for evaluating the performance of the filter. Instead, the experiments are performed on simulated range profiles. Artificially generated measurement noise was added to the range profiles so that they more closely resemble measured range profiles.

Finally, section 4.6 discusses the experimental results, followed by the conclusions in section 4.7.

4.2 Phase Response of a Point Scatterer

Section 2.3 defined a range profile as the squared amplitude of the complex-valued radar return $\mathbf{g}$. This transformation discards all phase information in $\mathbf{g}$, so only information on the distribution of scattering centres in slant range is retained. However, the phase of $\mathbf{g}$ does contain information.

Section 2.3 showed that the elements g_l of the radar return of a point scatterer, located at some position $\mathbf{x}$ in the aircraft-fixed coordinate system and observed at aspect angle $\boldsymbol{\theta}$, are given by

$$g_l = A\left(\frac{\sin \pi l'}{\sin \frac{\pi l'}{L+1}}\right) \exp i\phi_l(\boldsymbol{\theta}, \mathbf{x}), \tag{4.1}$$

$$\phi_l(\boldsymbol{\theta}, \mathbf{x}) = \phi_l^0 - \frac{4\pi \bar{f}}{c} r(\boldsymbol{\theta}, \mathbf{x}), \tag{4.2}$$

where l', $\bar{f}$ and ϕ_l are defined as

$$l' = l - \frac{2B}{c} r(\boldsymbol{\theta}, \mathbf{x}), \tag{4.3}$$

$$\bar{f} = f_0 + \frac{1}{2} L \Delta f, \tag{4.4}$$

$$\phi_l^0 = \frac{\pi L}{L+1} l. \tag{4.5}$$

Note that the phase $\phi_l(\boldsymbol{\theta}, \mathbf{x})$ in equation (4.2) is defined relative to the phase of the origin of the aircraft-fixed coordinate system, which is assumed to be zero.

Equation (2.6) defined a linear approximation of slant range as a function of aspect angle as

$$r(\boldsymbol{\theta} + \Delta\boldsymbol{\theta}, \mathbf{x}) \approx r(\boldsymbol{\theta}, \mathbf{x}) + r_\perp^a \, \Delta\theta^a + r_\perp^e \, \Delta\theta^e, \tag{4.6}$$

where the *cross ranges* $r_\perp^a$ and $r_\perp^e$ are defined in (2.7). Combining (4.2) and (4.6) shows that the derivative of the phase to θ^a and θ^e is respectively given by

$$\frac{d}{d\theta^a} \phi_l(\boldsymbol{\theta}, \mathbf{x}) = -\frac{4\pi \bar{f}}{c} r_\perp^a, \tag{4.7}$$

$$\frac{d}{d\theta^e} \phi_l(\boldsymbol{\theta}, \mathbf{x}) = -\frac{4\pi \bar{f}}{c} r_\perp^e. \tag{4.8}$$

Note that (4.7) and (4.8) only hold when the point scatterer remains within the range bin l. If rotational range migration causes the scatterer to move from range bin l to range bin $l+1$, a phase jump of size $\pi L/(L+1)$ occurs due to (4.5).

For the remainder of this chapter it is assumed that the direction of the rotation vector of the aircraft is constant, or equivalently, that the plane of rotation is constant. Let α denote the angle of rotation in this plane with respect to an initial aspect angle $\boldsymbol{\theta}_0$. The corresponding *generic cross range* $r_\perp$ is defined as

$$r_\perp = -\frac{c}{4\pi \bar{f}} \frac{d}{d\alpha} \phi_l(\alpha), \tag{4.9}$$

and so a linear approximation around $\boldsymbol{\theta}_0$ of the phase as a function of the rotation α is given by

$$\phi_l(\alpha) = \phi_l(\boldsymbol{\theta}_0) - \frac{4\pi \bar{f}}{c} r_\perp \alpha. \tag{4.10}$$

Equation (4.10) provides the theoretical basis for the filter for aicraft rotations, which will be derived in the next section.

4.3 Filter Overview

Chapter 2 defined the concept of a leg: a sequence of N successive HRR measurements obtained by repeatedly transmitting the full waveform. Let $\mathbf{g}_n$, $n = 1, \ldots, N$ denote the measured radar returns, i.e. the Fourier transforms of the coherent responses after windowing and zero-padding. The corresponding range profiles are denoted by $\mathbf{y}_n$.

During the measurement of the leg, the aspect angle at which the aircraft is observed will in general change due to motions of the aircraft. If the total duration of the measurement is short, it can be assumed that the aircraft plane of rotation is constant. Let α_n, denote the angle of rotation of the aircraft in this plane, measured with respect to some initial aspect angle $\boldsymbol{\theta}_0$.

For the moment assume that the radar returns $\mathbf{g}_n$ contain the contributions of a total of M point scatterers. Let ϕ_{nm} denote the phase of the mth point scatterer in the nth profile. It is assumed that the measurements ϕ_{nm} are corrected for the effects of rotational range migration, and so the subscript l is not used from now on. Details of the signal processing involved in extracting phase measurements from the leg and correcting the phases are provided in section 4.4. Under the stated assumptions it follows from equation (4.10) that, for small values of α, the phases ϕ_{nm} are related to the rotations α_n by

$$\phi_{nm} = \phi_m(\boldsymbol{\theta}_0) + \rho_m \alpha_n, \tag{4.11}$$

where $\rho_m = -4\pi \bar{f} {r_\perp}^m / c$, and ${r_\perp}^m$ is the generic cross range of the mth scatterer as defined in (4.9).

Both the rotations α_n and the phases ϕ_{nm} are subject to measurement noise, and so the linear relationship (4.11) will not hold exactly. The remainder of this section defines a filter for both the rotation and phase measurements, which finds new estimates of both the rotations α_n and the phases ϕ_{nm} such that their linear dependency holds as close as possible.

Let σ_α^2 and σ_ϕ^2 denote the error on rotation and phase measurements respectively. The filter provides new estimates $\bar{\alpha}_n$ and $\bar{\phi}_{nm}$, which minimise a chi-square function defined as

$$\chi^2 = \sum_n \left(\frac{(\alpha_n - \bar{\alpha}_n)^2}{\sigma_\alpha^2} + \sum_m \frac{(\phi_{nm} - \bar{\phi}_{nm})^2}{\sigma_\phi^2} \right), \tag{4.12}$$

subject to the linearity condition (4.11).

The filter operates as follows. First, the rotations α_n and phases ϕ_{nm} are used to construct a design matrix $\mathbf{A}$,

$$\mathbf{A} = \begin{bmatrix} \alpha_1/\sigma_\alpha & \alpha_2/\sigma_\alpha & \ldots & \alpha_N/\sigma_\alpha \\ \phi_{11}/\sigma_\phi & \phi_{21}/\sigma_\phi & \ldots & \phi_{N1}/\sigma_\phi \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{1M}/\sigma_\phi & \phi_{2M}/\sigma_\phi & \ldots & \phi_{NM}/\sigma_\phi \end{bmatrix}. \tag{4.13}$$

Then the column mean is subtracted from $\mathbf{A}$ to construct a zero-mean matrix denoted by $\mathbf{A}_0$. The columns of $\mathbf{A}_0$ are not independent, but instead are all related through (4.11). The matrix columns, when interpreted as vectors in a $M + 1$ dimensional vector space, span a subspace in that vector space. If no measurement noise were present, $\mathbf{A}_0$ would be of rank one, and the subspace spanned by the columns of $\mathbf{A}_0$ would be one-dimensional (i.e. a line). However, the measurement error on both α_n and ϕ_{nm} causes the rank of $\mathbf{A}_0$ to be larger than one. The task of the filter is to approximate $\mathbf{A}_0$ with a matrix $\bar{\mathbf{A}}_0$ of exactly rank one.

The optimal approximation $\bar{\mathbf{A}}_0$ of rank one of $\mathbf{A}_0$ can be found using the Singular Value Decomposition (SVD) [44] as follows. First, the SVD of $\mathbf{A}_0$ is computed,

$$\mathbf{A}_0 = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathrm{T}}. \tag{4.14}$$

The matrix $\mathbf{S}$ is diagonal, and the elements on the diagonal are the singular values λ_i of $\mathbf{A}$. The matrices $\mathbf{U}$ and $\mathbf{V}$ are both orthogonal in the sense that their columns $\mathbf{u}_i$ and $\mathbf{v}_j$ are orthonormal,

$$\mathbf{u}_i \cdot \mathbf{u}_j = \delta_{ij}, \tag{4.15}$$

$$\mathbf{v}_i \cdot \mathbf{v}_j = \delta_{ij}. \tag{4.16}$$

From the matrices $\mathbf{U}$, $\mathbf{S}$ and $\mathbf{V}$ the elements of the rank one matrix $\bar{\mathbf{A}}_0$ are defined as

$$(\bar{\mathbf{A}}_0)_{nm} = \lambda_1 u_{n1} v_{1m}. \tag{4.17}$$

Finally, the column mean of $\mathbf{A}$ is added to each column of $\bar{\mathbf{A}}_0$ to construct the matrix $\bar{\mathbf{A}}$

$$\bar{\mathbf{A}} = \begin{bmatrix} \bar{\alpha}_1/\sigma_\alpha & \bar{\alpha}_2/\sigma_\alpha & \dots & \bar{\alpha}_N/\sigma_\alpha \\ \bar{\phi}_{11}/\sigma_\phi & \bar{\phi}_{21}/\sigma_\phi & \dots & \bar{\phi}_{N1}/\sigma_\phi \\ \vdots & \vdots & \ddots & \vdots \\ \bar{\phi}_{1M}/\sigma_\phi & \bar{\phi}_{2M}/\sigma_\phi & \dots & \bar{\phi}_{NM}/\sigma_\phi \end{bmatrix}, \tag{4.18}$$

whose elements are the projections of the columns of $\mathbf{A}$ on the one-dimensional subspace spanned by the columns of $\mathbf{A}_0$. It can be shown that the resulting elements of $\bar{\mathbf{A}}$ minimise the chi-squared function defined in (4.12).

4.4 Feature Extraction

The previous section described a filter for combined rotation and phase measurements. This section describes the procedure used for extracting the necessary phase measurements ϕ_{nm} from a leg of radar returns $\mathbf{g}_n$, $n = 1, \dots, N$.

The radar returns contain the contribution of several distinct scatterers, which show up as amplitude peaks of $\mathbf{g}_n$, or equivalently of the corresponding range profiles $\mathbf{y}_n$ (see chapter 2). Section 4.4.1 describes a method for extracting peak locations and corresponding phases from the radar returns $\mathbf{g}_n$.

The next task of the feature extraction process is to decide which of the extracted peaks originate from a single scatterer. Figure 4.1(a) shows a leg of range profiles. In the figure several distinct tracks can be observed, where each track consists of a series of amplitude peaks caused by a single scatterer. Therefore, the task of assigning peaks to scatterers is essentially a tracking problem, where the features to track are the set of significant peaks and corresponding phases in the profiles in the leg. Section 4.4.2 describes the tracking algorithm used in this chapter.

The final task of the feature extraction process is to correct the phase measurements for the effects of both translational and rotational range migration, which is discussed in section 4.4.3.

4.4.1 Peak Detection

Amplitude peaks will be extracted from range profiles $\mathbf{y}_n$, whose range bins contain the squared magnitudes of the corresponding range bins in $\mathbf{g}_n$ (as discussed in section 2.3). First, all range bins containing local maxima of $\mathbf{y}_n$ are detected. Let K_n denote the total number of maxima in $\mathbf{y}_n$. For each maximum, its slant range r_{nk} and amplitude a_{nk}, $k = 1, \dots, K_n$ is estimated with sub-bin precision by fitting quadratic polynomial to the amplitudes of the range bin containing the maximum and the two adjoining range bins on both sides. Then, the same five range bins are used to estimate the phase ϕ_{nk} of $\mathbf{g}_n$ at slant range r_{nk} using a linear interpolation of the phases of the five range bins in $\mathbf{g}_n$. It is assumed that the range bins of $\mathbf{g}_n$ are aligned internally, for instance using maximum correlation as discussed in the previous chapter.

Due to measurement noise, some of the extracted peaks are not caused by scattering, but are caused by the noise process. Therefore, in the next step of the feature extraction process, the mean amplitude of the noise process will be estimated, and only peaks whose amplitude is significantly higher than the noise level will be retained.

The longest aircraft used in this thesis (the Boeing 747) is about 70 metres long, while the unambiguous range interval R_u is about 107 metres. Therefore, a section of at least 30 metres in each range profile $\mathbf{y}_n$ does not contain radar returns from the aircraft, i.e., it only contains noise. A simple heuristic to locate this section is to define it as that consecutive region of 20 metres length with the lowest average energy. The noise level μ is estimated as the mean amplitude in this section of $\mathbf{y}_n$.

Finally, only peaks with a signal-to-noise ratio $\mathrm{SNR}(a_{nk}, \mu)$, defined as

$$\mathrm{SNR}(a_{nk}, \mu) = 10 \log \frac{a_{nk}}{\mu}, \tag{4.19}$$

above an a priori defined threshold are retained and used for subsequent processing.

4.4.2 Peak Tracking

The next task is to decide which peaks (r_{nk}, a_{nk}) in the remaining set originate from the same scatterer. As discussed before, this is essentially a tracking problem.

Many different tracking algorithms have been reported in literature, each with its own advantages and disadvantages. For the work presented in this chapter a tracking algorithm originally reported in [49] is used. The main advantage of this tracking algorithm is that it by construction provides a one-to-one correspondence between peaks $\mathbf{r}_n$ and $\mathbf{r}_{n+1}$ in successive profiles.

The tracking algorithm establishes correspondences between the detected peaks in each possible pair of successive profiles n and $n+1$. It operates solely on the set of distances between the slant ranges r_{nk} and $r_{(n+1)k}$ of the peaks in two successive profiles – the amplitudes a_{nk} are not used for tracking.

At the start of the algorithm, for each pair of successive range profiles $\mathbf{y}_n$ and $\mathbf{y}_{n+1}$, a $K_n \times K_{n+1}$ *proximity matrix* $\mathbf{M}$ is constructed. The elements of $\mathbf{M}$ are defined as

$$m_{kl} = \exp -\frac{d_{kl}^2}{2\sigma_d^2}, \quad \text{for } k = 1, \ldots, K_n \text{ and } l = 1, \ldots, K_{n+1}, \tag{4.20}$$

where $d_{kl} = |r_{nk} - r_{(n+1)l}|$ is the distance in slant range between the peak locations r_{nk} and $r_{(n+1)l}$. The variance σ_d^2 is a free parameter of the algorithm.

The goal of the algorithm is to find a so-called *pairing matrix* $\mathbf{P}$, which minimises

$$\mathrm{trace}\, \mathbf{P}^T \mathbf{G} = \sum_k \sum_l p_{nm} m_{kl}. \tag{4.21}$$

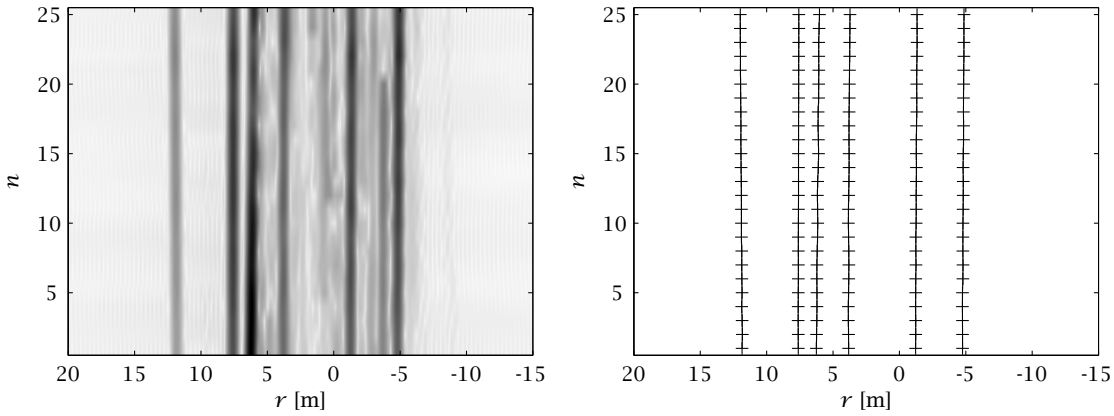
To construct the pairing matrix $\mathbf{P}$, first the singular value decomposition of $\mathbf{M}$,

$$\mathbf{M} = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathrm{T}}, \tag{4.22}$$

is calculated. Then, a diagonal matrix $\mathbf{S}'$ is constructed by replacing the diagonal elements of $\mathbf{S}$ with the value 1. In [49] it is shown that the paring matrix $\mathbf{P}$, defined by

$$\mathbf{P} = \mathbf{U}\mathbf{S}'\mathbf{V}, \tag{4.23}$$

minimizes (4.21).



(a) Example of leg, aligned using maximum correlation.

(b) Tracks detected from the leg.

Figure 4.1: Example of the output of the peak tracker.

The elements P_{kl} of $\mathbf{P}$ are a measure of the likelihood that the peak at r_{nk} and the peak at $r_{(n+1)l}$ belong to the same track. Therefore, if P_{kl} is both a row and a column maximum of $\mathbf{P}$, the peaks r_{nk} and $r_{(n+1)l}$ are assigned to the same track. Since the rows of $\mathbf{P}$ are orthonormal, the assignments of peaks to tracks is one-to-one. Figure 4.1(b) shows an example of the tracks found by this algorithm from the leg of range profiles shown in figure 4.1(a).

Although this tracking algorithm works well for our purposes, some post-processing of the established tracks is still required. First of all, to be useful for the filter described in the previous section, all tracks should be of length N, i.e., a track should be present in the complete leg. This is essentially a *missing data* problem. Although one could try to use interpolation techniques to artificially produce full-length tracks, in this chapter tracks of insufficient length are simply removed from the set.

Another problem is that sometimes large 'jumps' in slant range can be observed in a number of tracks. This is a results of the fact that the algorithm searches for one-to-one correspondences – after identifying the most probable matches, the last remaining possible match has to be assumed correct, no matter how unlikely. Therefore, tracks which exhibit slant range jumps larger than an a priori defined threshold are also removed from the set.

The final output of the complete tracking algorithm is a set of M tracks of peak locations r_{nm}, and corresponding (uncorrected) phases ϕ_{nm}

4.4.3 Phase Estimation

The final step of the feature extraction process is to correct the previously extracted phases ϕ_{nm} for the effects of translational and rotational range migration.

Section 4.2 assumed that the origin of the aircraft-fixed coordinate system is always observed in a range profile at zero slant range and with (by definition) zero phase. However, translational range migration causes the origin of the aircraft fixed coordinate system to appear at essentially random slant ranges in the different profiles in the leg. Although it has been assumed that the range bins of the $\mathbf{g}_n$ are aligned with respect to each other, the extracted phases still need to be corrected for the effects of TRM.

In section 2.2 it was noted that the origin of the aircraft-fixed coordinate system is fixed but

arbitrary. The simplest method for correcting the phase measurements for the effects of TRM is therefore to choose one single track, say $m = 1$, from the set of detected tracks, and to place the origin of the aircraft-fixed coordinate system at the position of the corresponding scatterer. Then, each phase measurement ϕ_{nm} can be corrected for TRM through

$$\phi_{nm} \rightarrow \phi_{nm} - \phi_{n1}, \tag{4.24}$$

$$r_{nm} \rightarrow r_{nm} - r_{n1}. \tag{4.25}$$

After this transformation, all phase measurements ϕ_{nm} are now relative to the origin of the aircraft-fixed coordinate system.

Next, the phase measurements need to be corrected for the effects of rotational range migration. Section 4.2 showed that a phase jump of $\pi L/(L+1)$ occurs if a point scatterer moves from range bin l to range bin $l+1$. Fortunately, rotational range migration can easily be detected from the slant range measurements r_{nm}, and so correcting the phase measurements ϕ_{nm} is trivial.

4.5 Experiments

This section describes the experiments performed to test the performance of the filter for aircraft rotations described in this chapter.

To evaluate the performance of the filter, i.e., the increase in quality of aircraft rotation estimates, the true values of the rotations need to be known. Unfortunately, all HRR measurements in the ORFEO measurement campaign were obtained from targets of opportunity, and so no onboard equipment was available to provide a ground truth on aspect angles and the corresponding rotations. Therefore, the ORFEO data can not be used for testing the filter. Instead, the filter will be tested on simulated profiles, obtained using the RAPPORT software package discussed in section 2.6.

Section 4.5.1 discusses the data set used for the experiments. To test the filter figure 2.4 under more or less realistic conditions, artificially generated noise will be added to both the simulated measurements and the true aspect angle. Section 4.5 discusses the actual experiments performed.

4.5.1 Data Description

The data used for the experiments consist of simulated legs of HRR measurements of the five aircraft shown in figure 2.4. Simulations were performed for each aircraft at the aspect angles shown in figure 4.2.

From the legs shown in figure 4.2 sub-legs containing 20 measurements each were constructed as follows. Let $c = 1, \ldots, C$ denote the class label of each aircraft. For each class c, the first sub-leg contained the first 20 measurements in the leg. The second and subsequent sub-legs contained the 19 last measurements of the previous sub-leg and the next measurement in the whole leg. The total number of sub-legs for each class is denoted by N^c.

Recall that a single HRR measurement consists of a set of coherent responses $\mathbf{G}$ measured in the frequency domain (see chapter 2). To simulate the effects of measurement noise, artificial complex-valued Gaussian noise to the coherent responses in each sub-leg. The standard deviation σ_G of the Gaussian noise is defined in terms of a signal-to-noise ratio defined as

$$\mathrm{SNR}(\sigma_G) = 20 \log \frac{|\bar{\mathbf{G}}|}{\sigma_G}, \tag{4.26}$$

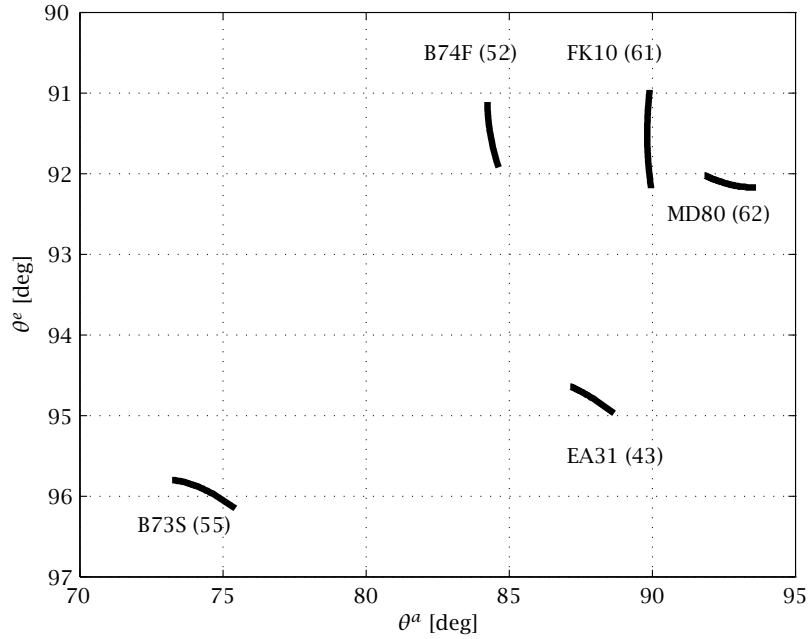


Figure 4.2: The aspect angles of HRR measurements in the test set for each class. The numbers following the aircraft type are the number of measurements in each leg.

where $|\bar{\mathbf{G}}|$ denotes the average absolute value of the elements of $\mathbf{G}$. After measurement noise has been added to the coherent responses, the effects of TRM were simulated by multiplying each set of coherent responses with a random phase shift.

Then, the resulting coherent responses were Hamming weighted and zero-padded to 1024 elements. Finally, an IDFT was performed to produce the radar returns $\mathbf{g}$ in the time domain.

For each sub-leg $n = 1,\ldots,N^c$ for each class c, the true rotations $\alpha_i^{c,n}$, $i = 1\ldots,10$ were calculated from the aspect angles $\boldsymbol{\theta}_n$. Rotations are defined with respect to the first aspect angle of the leg $\boldsymbol{\theta}_1$, and so $\alpha_1^{c,n} \equiv 0$. To simulate the error on rotation estimates present in actual HRR measurements. artificially generated Gaussian noise was added to the true rotations, resulting in rotation 'estimates' $\bar{\alpha}_i^{c,n}$. The standard deviation of the Gaussian noise was chosen to be $\sigma_\alpha = 2$ degrees, which corresponds to the estimated errors on rotation estimates for the ORFEO data.

4.5.2 Filter Performance Evaluation

From each sub-leg, phase measurements were extracted using the method described in section 4.4. Then, the filter described in section 4.3 was applied to the noisy rotation estimates $\hat{\alpha}_i^{c,n}$ and the extracted phases. The resulting updated rotation estimates are denoted by $\bar{\alpha}_i^{c,n}$.

Due to the noise added to the measurements, it is possible that no tracks can be established from the peaks extracted from the range profiles in the leg. In that case the algorithm has failed to provide improved rotation estimates, and the 'new' rotation estimates $\bar{\alpha}_i^{c,n}$ are chosen to be equal to the original noisy estimates $\hat{\alpha}_i^{c.n}$.

The *improvement factor* $\nu^{c,n}$ measures the performance of the filter for each sub-leg, and is defined as

$$\nu^{c,n} = \frac{\sum_{i=1}^{20}(\alpha_i^{c,n} - \bar{\alpha}_i^{c,n})^2}{\sum_{i=1}^{20}(\alpha_i^{c,n} - \hat{\alpha}_i^{c,n})^2}. \tag{4.27}$$

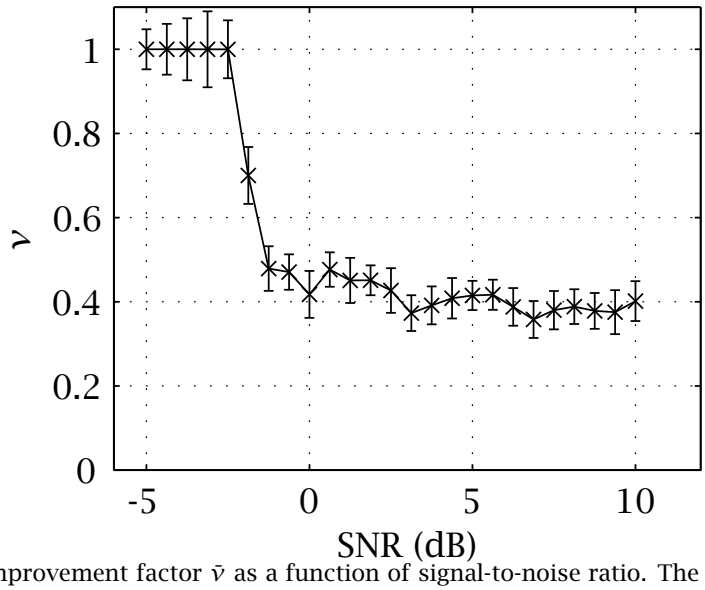


Figure 4.3: Mean improvement factor $\bar{\nu}$ as a function of signal-to-noise ratio. The bars denote the standard deviation σ_ν.

Note that $\nu^{c,n} = 1$ if the algorithm has failed, since in that case $\tilde{\alpha}_i^{c,n} = \hat{\alpha}_i^{c,n}$.

From the set of improvement factors $\nu_i^{c,n}$, the mean $\bar{\nu}$ and standard deviation σ_ν,

$$\bar{\nu} = \frac{1}{C} \sum_{c=1}^{C} \frac{1}{N^c} \sum_{n=1}^{N^c} \nu^{c,n}, \tag{4.28}$$

$$\sigma_\nu = \frac{1}{C} \sum_{c=1}^{C} \frac{1}{N^c} \sum_{n=1}^{N^c} (\nu^{c,n} - \bar{\nu})^2, \tag{4.29}$$

were determined. The value of $\bar{\nu}$ and σ_ν was measured several times for different values of the signal-to-noise ratio, ranging from -5 to 10 dB.

4.6 Results

Figure 4.3 shows $\bar{\nu}$ and σ_ν as a function of the signal-to-noise ratio. For low values of the signal-to-noise ratio no tracks can be established, and so the mean improvement factor is 1. For signal-to-noise ratios higher than roughly -2 dB, the algorithm is able to find enough tracks, and the improvement factor decreases until it reaches values of about 0.4, after which it remains constant.

4.7 Conclusions

This chapter introduced a method for increasing the accuracy of aircraft rotation estimates obtained from tracking data. On the basis of a point scatter model a filter was developed, which uses a combination of the original rotation estimates and phase information extracted from range profiles to update the rotation estimates.

To extract the necessary phase information from the range profiles an extensive feature extraction method was developed. First, tracks of amplitude peaks were detected using a simple

yet effective tracking algorithm originally described in [49]. From these tracks phase sequences were extracted, which were subsequently corrected for the effects of translational and rotational range migration.

Since no ground truth on aspect angles are available for the range profiles measured during the ORFEO campaign, the experiments were necessarily restricted to simulated profiles. However, the effects of measurement noise and translational range migration were artificially added to the simulated range profiles, and so the resulting data set is similar to actual HRR measurements.

To evaluate the performance of the filter, an improvement ratio ν was introduced. It has been shown that the filter does indeed improve aircraft rotation estimates (typically $\nu \approx 0.4$), provided that the signal-to-noise ratio is sufficiently high (larger than -2 dB). For smaller signal-to-noise ratios, the algorithm fails to provide more accurate rotation estimates, resulting in an improvement factor of 1.

Although these are promising results, a definitive answer on the applicability of this method on actual measured range profiles does require further testing.

5 A Generative Model of Point Scattering

5.1 Introduction

A drawback of using range profiles for classifying aircraft is that measuring sufficient range profiles from which a statistical classifier can be constructed is not feasible (see chapter 1). A solution to this problem is to use a training set containing simulated range profiles. Simulated range profiles differ significantly from measured range profiles (see section 2.6). If a classifier for measured profiles is to be trained from simulated profiles, care must be taken to ensure that the classifier bases its decision solely on the common features between simulated and measured profiles.

A reasonable assumption is that at least some of the dominant scatterers found in simulated profiles will also be present in measured profiles. For example, when observing a commercial airliner from approximately broadside, one would expect scatterers to be located at the location of jet engines and the body of the aircraft in both measured and simulated profiles. These dominant scatterers cause peaks at a certain location with a certain amplitude in both simulated and measured profiles. The peak locations, which are determined by the positions of the associated scatterers on the aircraft, should be identical (within reasonable bounds) in both cases.

This is in general not true for peak amplitudes. It is very difficult to predict the (relative) amplitude of peaks in simulated profiles, since not all scattering behaviour is included in the simulation. If one or more scatterers are not present in the simulated profile, the (relative) amplitude of the remaining peaks are useless for comparison with measured profiles.

This chapter therefore assumes that the common features between measured and simulated range profiles are the slant ranges of amplitude peaks.

The use of peak locations (in combination with amplitude information) as features for classification has previously been discussed in [8, 50], where a peak location probability function is defined which determines for each location in a profile the probability that a peak occurs there. This function is estimated by aggregating peak location information over all possible aspect angles. Since peak locations are dependent on aspect angle (through the effects of TRM), this procedure causes a 'widening' of the peak location distribution. This effect will be more severe for increasing numbers of poses, and so this approach can only be applied to training sets in which the pose does not change significantly.

This chapter presents a method which uses peak locations as features for classification, but does not build the classifier in the feature space. Instead, for each class the parameters of a *generative* model will be estimated, i.e., a set of parameters from which, for any given aspect angle, the statistical distribution of peak locations can be predicted. These generative models are then used to construct a maximum likelihood classifier for radar range profiles.

A simple and adequate model for the distribution of peak locations in range profiles is a configuration of point scatterers, including their visibility to model occlusions. The advantage of such a simple model is that its parameters (point scatter locations and visibility areas) can be automatically estimated from a training set of simulated range profiles. Once generative models for each aircraft are available, they can be used to determine the likelihood of an unknown series of profiles, conditioned on aircraft class.

In recent literature [29, 30], a more sophisticated parametric model of radar scattering has been developed. This model includes more complicated scattering, and contains the point

scatterer as a special case. This chapter uses a simpler model in order to limit the complexity of the parameter estimation algorithm.

Expressing the likelihood of peak locations given the model parameters requires the use of a set of *hidden variables*. The hidden variables of this problem are the assignments of observed peaks in the profiles to scatterers in the model. The optimal model parameters are estimated using an Expectation-Maximisation algorithm [51].

A similar parameter estimation problem is presented in [52], where the goal is to reconstruct the three-dimensional shape of an object from a set of images taken from different camera positions. The hidden variables of this problem describe which observed image features correspond to which 3D feature of the object. While the parameter estimation procedure presented in [52] provides the basis for the procedure presented in this chapter, there are some significant differences.

Section 5.2 provides detailed descriptions of the estimation of the parameters of the generative models. Section 5.3 shows how the estimated models can be used to construct a maximum likelihood classifier for range profiles. Sections 5.4 and 5.5 describe the experiments performed to test the models. Finally, section 5.6 contains conclusions based on the experimental results.

5.2 Estimation of Model Parameters

This section describes the construction of a model-based classifier for range profiles. The generative model used by the classifier is capable of predicting the distribution of amplitude peaks occurring in range profiles, given aircraft type and aspect angle. The parameters of the model are estimated from peak locations extracted from a training set containing simulated range profiles.

All significant peaks in simulated range profiles are assumed to be caused by point scatterers. There are a number of advantages in using such a simple model. The parameters of a point scatter model are easy to estimate given only measurements of peak locations in a set of profiles. Also, given such a model it is possible to predict peak locations for any given aspect angle, i.e., it allows us to interpolate between different aircraft poses. This is a crucial requirement for a classifier for range profiles.

5.2.1 Observed Variables and Model Parameters

The training set for each aircraft consists of N simulated range profiles. The aspect angles at which these profiles were collected are denoted by $\boldsymbol{\Theta} = \{\boldsymbol{\theta}_n\}$ for $n = 1, \ldots, N$.

The parameters of the point scatter models will be estimated solely from the positions of local amplitude maxima, i.e., *peaks*, in each profile. As explained in chapter 2, range profiles are measurements of an aircraft's radar reflectivity as a function of slant range r, defined in section 2.2. The positions of amplitude peaks can therefore be expressed as a set of slant ranges r_{nk}, where n is the index of the profile in the training set, and $k = 1, \ldots, K_n$, where K_n is the total number of amplitude maxima in the nth profile. Note that in general different profiles will contain a different number of amplitude maxima, and so K_n will be different for each profile. The complete collection of amplitude peak positions $\mathbf{R} = \{r_{nk}\}$ are the *observed variables* of the problem.

The distribution of peak locations is modeled by a generative point scatter model containing M point scatterers located at positions $\mathbf{x}_m = (x_m, y_m, z_m)$, $m = 1, \ldots, M$, in the aircraft-fixed coordinate system defined in section 2.2. For each scatterer, the corresponding peak location distribution, when observed at aspect angle $\boldsymbol{\theta}$, is modeled as a Gaussian distribution with mean

$r(\boldsymbol{\theta}, \mathbf{x}_m)$ (defined in equation (2.5)) and variance σ_m^2,

$$p(r|\mathbf{x}_m, \sigma_m^2) = \mathcal{N}(r; r(\boldsymbol{\theta}, \mathbf{x}_m), \sigma_m^2). \tag{5.1}$$

Note that the variances σ_m^2 are parameters of the generative models, and are assumed to be independent of aspect angle.

Section 2.4 introduced the main sources of range profile variability. The simulated range profiles in the training set are not influenced by translational range migration or measurement noise, but they are influenced by speckle and occlusion.

Both speckle and occlusion have an effect on the *visibility* of a scatterer. Speckle can cause the return of several scatterers present in a given range bin to cancel out completely, and as a consequence no amplitude peak will be present in that range bin. Similarly, if a scatterer is occluded at a given aspect angle, it does not generate an amplitude peak in the corresponding profile.

The effects of both speckle and occlusion will be collectively modelled by introducing a *visibility matrix* $\mathbf{V}$. It is a $N \times M$ matrix with elements v_{nm}, with v_{nm} defined as the probability of scatterer m being visible when observed at aspect angle $\boldsymbol{\theta}_n$. If $v_{nm} = 1$, the scatterer is visible, while a value of $v_{nm} = 0$ means it is occluded. Values of v_{nm} between 0 and 1 indicate that speckle is causing the visibility to rapidly fluctuate over a small changes of aspect angle (see also section 5.2.6).

For notational convenience, $\boldsymbol{\Psi} = \{\mathbf{X}, \boldsymbol{\sigma}, \mathbf{V}\}$ is used to denote the complete set of model parameters $\mathbf{X} = \{\mathbf{x}_m\}$, $\boldsymbol{\sigma} = \{\boldsymbol{\sigma}_m\}$ and $\mathbf{V} = \{v_{nm}\}$.

5.2.2 Hidden Variables

The optimal model parameters $\boldsymbol{\Psi}^*$ are found by maximising the joint log likelihood of the observed variables $\mathbf{R}$ as a function of the model parameters $\boldsymbol{\Psi}$, given the aspect angles $\boldsymbol{\Theta}$,

$$\boldsymbol{\Psi}^* = \arg\max_{\boldsymbol{\Psi}} \log p(\mathbf{R}|\boldsymbol{\Psi}; \boldsymbol{\Theta}). \tag{5.2}$$

The likelihood of observing the peak positions $\mathbf{R}$ is dependent on the way scatterers in the model are *assigned* to the peaks, i.e, to solve (5.2) it is necessary to know which of the observed peaks r_{nk} are measurements of the slant range positions of which of the scatterers $\mathbf{x}_m$. An example of an assignment is shown in figure 5.1.

Recall that in profile n there are a total of K_n detected peaks, which have to be assigned to a total of M scatterers. In general $K_n \neq M$, both because peaks may be present in the profile which were not caused by any of the scatterers, and because of any of the peak detection issues described in section 2.4.

Assignments are described by a set of assignment variables $\boldsymbol{\Omega}$. Let ω_{nk} denote an assignment between a detected peak and a scatterer,

$$\omega_{nk} \in \{0, 1, \ldots, M\}, \tag{5.3}$$

The statement $\omega_{nk} = m$ means that the measured slant range location r_{nk} of the kth peak in the nth profile is a measurement of $\hat{r}_{nm} = \hat{r}(\boldsymbol{\theta}_n, \mathbf{x}_m)$, while $\omega_{nk} = 0$ means this peak is not assigned to any scatterer. An overview of the relationship between scatterers, peaks and assignments is shown in figure 5.1.

The set of assignments between peaks and scatterers in the nth profile are denoted by $\boldsymbol{\omega}_n = \{\omega_{nk}\}$, and the complete set of assignments between peaks and scatterers over all profiles by $\boldsymbol{\Omega} = \{\boldsymbol{\omega}_n\}$.

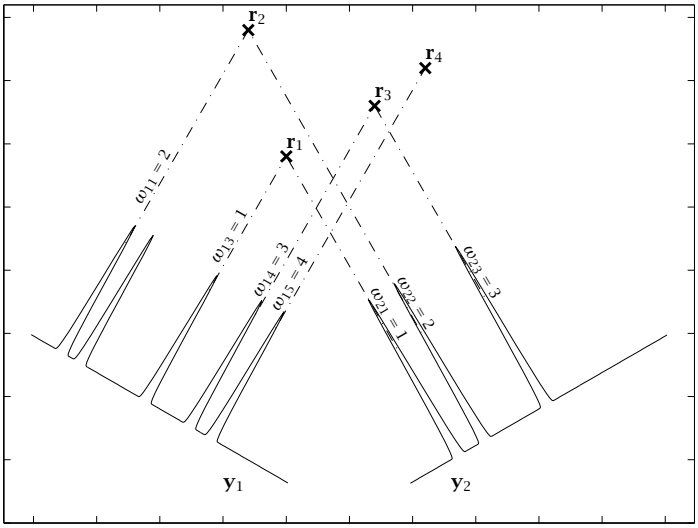


Figure 5.1: Example of possible assignments ω_{nk} between 4 scatterers $\mathbf{x}_m$ (here shown in 2 dimensions instead of 3), and peaks in two profiles $\mathbf{z}_1$ and $\mathbf{z}_2$. Profile $\mathbf{z}_1$, measured at aspect angle $\boldsymbol{\theta}_1$, contains 5 peaks at slant ranges r_{1k}, where $k = 1, \ldots, 5$. The second profile, $\mathbf{z}_2$, measured at a different aspect angle $\boldsymbol{\theta}_2$, contains only 3 peaks at slant ranges r_{2k}, where now $k = 1, \ldots, 3$.

Not all possible assignment are *valid* assignments. An assignment $\boldsymbol{\omega}_n$ is valid if each peak is assigned to at most one scatterer, and each scatterer is assigned to at most one peak. The first condition is always met due to the way assignments are encoded. The second condition must be enforced through other means, which are described later in this chapter.

Although the assignments $\boldsymbol{\Omega}$ are required to be able to express the likelihood of the observations given the model, their correct value is unknown. They are the *hidden variables* of the problem. The joined log likelihood occurring in (5.2) can only be expressed as a *marginalisation* over the hidden variables $\boldsymbol{\Omega}$:

$$\log p(\mathbf{R}|\boldsymbol{\Psi};\boldsymbol{\Theta}) = \log \sum_{\boldsymbol{\Omega}} p(\mathbf{R},\boldsymbol{\Omega}|\boldsymbol{\Psi};\boldsymbol{\Theta}). \tag{5.4}$$

In section 5.2.4 an exact definition of this likelihood will be given.

5.2.3 Expectation-Maximisation Optimisation

From (5.2) and (5.4) it follows that the optimal model parameters $\boldsymbol{\Psi}^*$ are those that maximise the joint log likelihood of peak locations $\mathbf{R}$ and the hidden assignment variables $\boldsymbol{\Omega}$,

$$\boldsymbol{\Psi}^* = \arg\max_{\boldsymbol{\Psi}} \log \sum_{\boldsymbol{\Omega}} p(\mathbf{R},\boldsymbol{\Omega}|\boldsymbol{\Psi};\boldsymbol{\Theta}). \tag{5.5}$$

Due to the dependence of the log likelihood on the unknown assignments $\boldsymbol{\Omega}$, equation (5.5) can not be solved directly.

A common approach for maximum likelihood parameter estimation in the presence of hidden variables is to use an Expectation-Maximisation (EM)[51, 53, 54, 55, 56]. The main idea behind EM optimisation is that although the true values of the hidden variables are unknown, it might be possible to estimate a distribution over $\boldsymbol{\Omega}$, given the input data and an estimate of the model parameters. This distribution is then used to obtain an improved estimate of the model

parameters.

EM is an iterative algorithm, and consists of an *expectation* and a *maximisation* step. Let t be the number of the current iteration.

- In the expectation step, a distribution $q^t(\mathbf{\Omega})$ is computed such that

$$q^t(\mathbf{\Omega}) = p(\mathbf{\Omega}|\mathbf{R}, \mathbf{\Psi}^t; \mathbf{\Theta}), \tag{5.6}$$

  where $\mathbf{\Psi}^t$ is the current best estimate of the model parameters.

- In the maximisation step, a new best estimate $\mathbf{\Psi}^{t+1}$ of $\mathbf{\Psi}^*$ is constructed by maximising the *expected* joined log likelihood of the observed and hidden variables, where the expectation value is taken with respect to $q^t(\mathbf{\Omega})$:

$$\mathbf{\Psi}^{t+1} = \arg\max_{\mathbf{\Psi}} E_{q^t}\left[\log p(\mathbf{R}, \mathbf{\Omega}|\mathbf{\Psi}; \mathbf{\Theta})\right]. \tag{5.7}$$

In appendix A it is shown that this procedure maximises $\log p(\mathbf{R}|\mathbf{\Psi}; \mathbf{\Theta})$. Detailed descriptions of the implementation of the expectation and maximisation steps are given in sections 5.2.5 and 5.2.6.

5.2.4 Derivation of Likelihoods

This section derives explicit expressions for the likelihoods occurring in equations (5.6) and (5.7).

The joint likelihood of the observed and hidden variables can be written as

$$p(\mathbf{R}, \mathbf{\Omega}|\mathbf{\Psi}; \mathbf{\Theta}) = p(\mathbf{R}|\mathbf{\Omega}, \mathbf{\Psi}; \mathbf{\Theta})\, p(\mathbf{\Omega}|\mathbf{\Psi}; \mathbf{\Theta}). \tag{5.8}$$

Since $p(\mathbf{R}|\mathbf{\Omega}, \mathbf{\Psi}; \mathbf{\Theta})$ is conditioned on the assignments $\mathbf{\Omega}$, this distribution is independent of the visibility matrix $\mathbf{V}$, since a scatterer assigned to any given peak is visible by definition. Likewise, a scatterer which has not been assigned to any peak invisible by definition. Therefore,

$$p(\mathbf{R}|\mathbf{\Omega}, \mathbf{\Psi}; \mathbf{\Theta}) = p(\mathbf{R}|\mathbf{\Omega}, \mathbf{X}, \boldsymbol{\sigma}; \mathbf{\Theta}). \tag{5.9}$$

The peak positions are assumed to be *conditionally* independent,

$$p(\mathbf{R}|\mathbf{\Omega}, \mathbf{X}, \boldsymbol{\sigma}; \mathbf{\Theta}) = \prod_{n=1}^{N} \prod_{k=1}^{K_n} p(r_{nk}|\omega_{nk}, \mathbf{X}, \boldsymbol{\sigma}; \boldsymbol{\theta}_n). \tag{5.10}$$

The distribution of the slant range r_{nk} of a peak assigned to scatterer m, i.e., $\omega_{nk} = m$, is defined as the Gaussian distribution given by

$$p(r_{nk}|\omega_{nk} = m, \mathbf{X}, \boldsymbol{\sigma}; \boldsymbol{\theta}_n) = \mathcal{N}(r_{nk}; r(\boldsymbol{\theta}_n, \mathbf{x}_m), \sigma_m^2). \tag{5.11}$$

If peak r_{nk} is *not* assigned to any scatterer, i.e., $\omega_{nk} = 0$, all slant range are equally likely, and so

$$p(r_{nk}|\omega_{nk} = 0, \mathbf{X}, \boldsymbol{\sigma}; \boldsymbol{\theta_n}) = \frac{1}{L}, \tag{5.12}$$

where L is the unambiguous range interval, i.e. the total length of the profile, defined in 2.2.

The second term in (5.8) is the *a priori* likelihood of the set of assignments $\boldsymbol{\Omega}$, given the model parameters $\boldsymbol{\Psi}$. It is assumed to be only dependent on the visibility matrix $\mathbf{V}$,

$$p(\boldsymbol{\Omega}|\boldsymbol{\Psi};\boldsymbol{\Theta}) = p(\boldsymbol{\Omega}|\mathbf{V}). \tag{5.13}$$

The a priori likelihood of assigning a scatterer m to *any* of the peaks in the nth profile is just v_{nm}, which is the probability of visibility of the mth scatterer in the nth profile. Assigning a scatterer which is considered occluded ($v_{nm} \approx 0$) is therefore considered very unlikely, while assigning a scatterer which is known to be visible ($v_{nm} \approx 1$) is considered very likely. Likewise, the a priori likelihood of *not* assigning the mth scatterer to a peak in the nth profile is $1 - v_{nm}$.

Let ϵ_{nm} be a binary variable indicating whether the mth scatterer has been assigned to *any* peak in the nth profile,

$$\epsilon_{nm} = \begin{cases} 1 & \text{if } m \in \boldsymbol{\omega}_n \\ 0 & \text{if } m \notin \boldsymbol{\omega}_n \end{cases}. \tag{5.14}$$

The a priori likelihood of an assignment $\boldsymbol{\Omega}$ is then given by

$$\begin{aligned} p(\boldsymbol{\Omega}|\mathbf{V}) &= \prod_{n=1}^{N} p(\boldsymbol{\omega}_n|\mathbf{v}_n) \\ &= \prod_{n=1}^{N} \prod_{m=1}^{M} (v_{nm})^{\epsilon_{nm}} (1 - v_{nm})^{1-\epsilon_{nm}}. \end{aligned} \tag{5.15}$$

Finally, the distribution $q^t(\boldsymbol{\Omega})$ defined in (5.6) can be rewritten using Bayes' rule as

$$\begin{aligned} q^t(\boldsymbol{\Omega}) &= p(\boldsymbol{\Omega}|\mathbf{R}, \boldsymbol{\Psi}^t; \boldsymbol{\Theta}) \\ &\propto p(\mathbf{R}|\boldsymbol{\Omega}, \boldsymbol{\Psi}^t; \boldsymbol{\Theta})\, p(\boldsymbol{\Omega}|\boldsymbol{\Psi}^t; \boldsymbol{\Theta}). \end{aligned} \tag{5.16}$$

Using the independence assumptions (5.11) and (5.15), equation (5.16) can be written as

$$q^t(\boldsymbol{\Omega}) \propto \prod_{n=1}^{N} p(\mathbf{r}_n|\boldsymbol{\omega}_n, \mathbf{X}^t, \boldsymbol{\sigma}^t; \boldsymbol{\theta}_n)\, p(\boldsymbol{\omega}_n|\mathbf{V}^t). \tag{5.17}$$

Since the space of possible assignments is discrete, $q^t(\boldsymbol{\Omega})$ is completely defined by a set of probabilities q^t_{nkm}, defined as

$$q^t_{nkm} = p(\omega_{nk} = m|\mathbf{R}, \boldsymbol{\Psi}^t; \boldsymbol{\Theta}), \tag{5.18}$$

i.e., q^t_{nkm} is the posterior likelihood of assigning the kth peak in the nth profile to scatterer m.

5.2.5 Description of the Expectation Step

Unfortunately, because of the constraints on valid assignments $\boldsymbol{\omega}_n$ (described in section 5.2.2), it is impossible to directly calculate (5.18). Instead, the q^t_{nkm} will be estimated by sampling the posterior distribution over valid assignments. From this set of samples the q^t_{nkm} is estimated as

$$q^t_{nkm} = \frac{1}{S} \sum_{s=1}^{S} \delta(\omega^s_{nk}, m), \tag{5.19}$$

where δ is the Kronecker delta function and S is the total number of samples.

A Markov chain Monte Carlo (MCMC) algorithm is used to obtain the samples. Monte Carlo

algorithms (or more specifically, Metropolis-Hastings algorithms [57, 58]) sample a target distribution by sequentially generating new samples from a *proposal distribution*, which are then either accepted or rejected based on the value of some acceptance ratio a.

In MCMC algorithms, the proposal distribution takes the form of a Markov process, which means that a new sample $\boldsymbol{\omega}'_n$ is generated probabilistically from the last accepted sample $\boldsymbol{\omega}_n$, according to a so-called *Markov transition probability* $p(\boldsymbol{\omega}'_n|\boldsymbol{\omega}_n)$.

The exact sampling process, which is performed for each profile separately, is as follows:

1. Start with a valid assignment $\boldsymbol{\omega}_n^s$, $s = 1$.

2. For $s = 1$ to S,

   a) Randomly generate a new valid assignment $\boldsymbol{\omega}'_n$ from $\boldsymbol{\omega}_n^s$ according to the Markov transition probability $p(\boldsymbol{\omega}'_n|\boldsymbol{\omega}_n^s)$.

   b) Calculate the acceptance ratio a,

$$a = \frac{p(\boldsymbol{\omega}'_n|\mathbf{r}_n, \boldsymbol{\Psi}^t; \boldsymbol{\theta}_n)}{p(\boldsymbol{\omega}_n^s|\mathbf{r}_n, \boldsymbol{\Psi}^t; \boldsymbol{\theta}_n)} \frac{p(\boldsymbol{\omega}_n^s|\boldsymbol{\omega}'_n)}{p(\boldsymbol{\omega}'_n|\boldsymbol{\omega}_n^s)}. \tag{5.20}$$

   c) If $a > 1$, accept $\boldsymbol{\omega}'_n$ by setting $\boldsymbol{\omega}_n^{s+1} = \boldsymbol{\omega}'_n$. Otherwise, accept $\boldsymbol{\omega}'_n$ with a probability a. If the proposal is rejected, set $\boldsymbol{\omega}_n^{s+1} = \boldsymbol{\omega}_n^s$.

The acceptance/rejection mechanism ensures that the final set of samples are drawn from $p(\boldsymbol{\Omega}|\mathbf{R}, \boldsymbol{\Psi}^t; \boldsymbol{\Theta})$. This property of the sampling process is, apart from some general restrictions [54], independent of the exact choice for the Markov transition probabilities. The Markov transition probabilities do however determine the efficiency of the sampling process. To ensure the algorithm explores the search space sufficiently while still generating high probability samples, the following method for generating new samples was used:

1. Randomly choose a scatterer $m \in 1, \ldots, M$.

2. Find the current assignment k of m, where $k \in 1, \ldots, K_n$ means m is currently assigned to peak k, i.e, $\omega_{nk}^s = m$, and $k = 0$ means m is considered occluded in the current value of ω_{nk}^s.

3. Randomly change the assignment for m from k to k', according to the probability $P(k'|k, m)$ defined as

$$P(k'|k, m) = \frac{p(k'|k, m)}{\sum_{k'} p(k'|k, m)},$$
$$p(k'|k, m) = \begin{cases} 0 & \text{if } k' = k, \\ (1 - v_{nm}) & \text{if } k' = 0 \text{ and } k' \neq k, \\ p(r_{nk'}|\mathbf{x}_m, \sigma_m; \boldsymbol{\theta}_n) \times v_{nm} & \text{if } k' \in 1, \ldots, K_n \text{ and } k' \neq k. \end{cases} \tag{5.21}$$

4. If m is now assigned to a peak formerly assigned to another scatterer m', repeat the above process for scatterer m'. Otherwise, stop.

The last step in the algorithm ensures that the resulting proposal assignment $\boldsymbol{\omega}'$ is always a valid assignment.

Finally, estimates of q_{nkm}^t are constructed from the resulting set of samples using equation (5.19). These estimates are then used to construct the expected joint log likelihood in the maximisation step.

5.2.6 Description of the Maximisation Step

In the maximisation step, a new best estimate $\mathbf{\Psi}^{t+1}$ of the model parameters is found by maximising the expected joint log likelihood of the observed and hidden variables,

$$\begin{aligned}\mathbf{\Psi}^{t+1} &= \arg\max_{\mathbf{\Psi}} E_{q^t}\left[\log p(\mathbf{R}, \mathbf{\Omega}|\mathbf{\Psi}; \Theta)\right] \\ &= \arg\max_{\mathbf{X},\boldsymbol{\sigma},\mathbf{V}} E_{q^t}\left[\log p(\mathbf{R}|\mathbf{\Omega}, \mathbf{X}, \boldsymbol{\sigma}; \Theta) + \log p(\mathbf{\Omega}|\mathbf{V})\right],\end{aligned} \tag{5.22}$$

where the last step uses the definitions presented in section 5.2.4. The expected joint log likelihood can be expressed in terms of the q^t_{nkm} (defined in equation (5.18)) as

$$E_{q^t}\left[\log p(\mathbf{R}|\mathbf{\Omega}, \mathbf{X}, \boldsymbol{\sigma}; \Theta)\right] = \sum_{n=1}^{N}\sum_{k=1}^{K}\sum_{m=0}^{M} q^t_{nkm} \log p(r_{nk}|w_{nk} = m, \mathbf{X}, \boldsymbol{\sigma}; \boldsymbol{\theta}_n). \tag{5.23}$$

The new best estimate $\mathbf{X}^{t+1}$ of $\mathbf{X}$ is most easily defined in terms of *virtual measurements* $\check{r}_{nm}$ and *virtual measurement variances* $\check{\sigma}^2_m$, defined as [52]

$$\begin{aligned}\check{r}_{nm} &\equiv \frac{\sum_{k=1}^{K_n} q^t_{nkm}\, r_{nk}}{\sum_{k=1}^{K_n} q^t_{nkm}}, \\ \check{\sigma}^2_m &\equiv \frac{\sigma^2_m}{\sum_{k=1}^{K_n} q^t_{nkm}},\end{aligned} \tag{5.24}$$

for $n = 1, \ldots, N$ and $m = 1, \ldots, M$. Using these definitions, equation (5.23) can be written as

$$E_{q^t}\left[\log p(\mathbf{R}|\mathbf{\Omega}, \mathbf{X}, \boldsymbol{\sigma}; \Theta)\right] = C + \sum_{n=1}^{N}\sum_{m=1}^{M} \frac{(\check{r}_{nm} - \hat{r}_{nm})^2}{2\check{\sigma}^2}. \tag{5.25}$$

where C is a constant which does not depend on $\mathbf{X}$. A new estimate $\mathbf{X}^{t+1}$ of the scatter positions can be found by maximising (5.25) using a General Least Squares solver.

Given the new position estimates $\mathbf{X}^{t+1}$, the maximum likelihood solution for the model variances are given by

$$(\sigma^{t+1}_m)^2 = \frac{\sum_{n=1}^{N}\sum_{k=1}^{K_n} q^t_{nkm}\,(r_{nk} - \hat{r}_{nm}))^2}{\sum_{n=1}^{N}\sum_{k=1}^{K_n} q^t_{nkm}}. \tag{5.26}$$

To estimate the visibility matrix, the expected log likelihood (5.15) can be written in terms of the assignments $\mathbf{\Omega}$ as

$$E_{q^t}\left[\log p(\mathbf{\Omega}|\mathbf{\Psi}; \Theta)\right] = \sum_n \sum_m \bar{\epsilon}_{nm} \log(v_{nm}) + (1 - \bar{\epsilon}_{nm}) \log(1 - v_{nm}), \tag{5.27}$$

where $\bar{\epsilon}_{nm}$ is defined as the expectation value of ϵ_{nm},

$$\bar{\epsilon}_{nm} = E_{q^t}\left[\epsilon_{nm}\right] = \sum_k q^t_{nkm}. \tag{5.28}$$

The maximum likelihood solution v^{t+1}_{nm} is the given by

$$v^{t+1}_{nm} = \bar{\epsilon}_{nm} = \sum_k q^t_{nkm}. \tag{5.29}$$

Equation (5.29) should be interpreted as follows: the probability of scatterer m being visible in profile n is equal to the sum over k of the probabilities that scatterer m is assigned to peak k.

Although the estimate (5.29) is the maximum likelihood solution, it does not reflect the physical properties of radar scattering correctly. The problem is that in the training set, a scatterer is either present in a profile, or it is not. Unconstrained optimisation of the visibility matrix using (5.29) would therefore result in a 'binary' visibility matrix, i.e., all elements v_{nm} would be either 0 or 1.

This is a problem in the case of speckle. As described in section 5.2, speckle causes a the visibility of a scatterer to vary rapidly over small changes in aspect angle. Therefore, speckle should be modelled as a probability of visibility between 0 and 1.

To remedy this problem, the visibility matrix is smoothed after the new estimates (5.29) have been calculated, by applying a Gaussian filter to the elements v_{nm}^{t+1},

$$v'^{t+1}_{nm} = \sum_{n'=1}^{N} \sum_{m'=1}^{M} W_{n'm'}\, v^{t+1}_{n'm'}, \tag{5.30}$$

$$W_{nm} = \frac{\left(2\pi\sigma_v^2\right)^{-1/2} \exp \frac{(\boldsymbol{\theta}_n - \boldsymbol{\theta}_{n'})^2}{2\sigma_v^2}}{\sum_{n'=1}^{N}\sum_{m'=1}^{M}\left(2\pi\sigma_v^2\right)^{-1/2} \exp \frac{(\boldsymbol{\theta}_n - \boldsymbol{\theta}_{n'})^2}{2\sigma_v^2}}. \tag{5.31}$$

The elements v'^{t+1}_{nm} form the new estimate of the visibility matrix $\mathbf{V}^{t+1}$.

A drawback of using this smoothing procedure is that the new estimate $\mathbf{V}^{t+1}$ no longer maximises the expected log likelihood (5.27). Consequently, the convergence of the parameter estimation algorithm is no longer guaranteed as it is for true EM algorithms. Nevertheless, the approach converges in practice, and the resulting estimate of $\mathbf{V}$ does more accurately reflect the physics of radar scattering.

5.3 Model-Based Classification

This section describes the construction of a maximum likelihood classifier for simulated range profiles. Although such a classifier has no practical application, it can validate the assumptions underlying the generative point scatter model if it performs well. Furthermore, the classifier described in this section is the basis for a maximum likelihood classifier for *measured* range profiles, described in chapter 6.

Let $c = 1, \ldots, C$ denote the class labels of the aircraft in the database. It is assumed that for each aircraft a generative point scatter model $\boldsymbol{\Psi}^c$ has been estimated. The task of the classifier is to assign a class label c^* to a leg of simulated profiles $\mathbf{Y} = \{\mathbf{y}_n\}$ with aspect angles $\boldsymbol{\Theta} = \{\boldsymbol{\theta}_n\}$.

First, the peak locations $\mathbf{R} = \{r_{nk}\}$ are extracted from the leg $\mathbf{Y}$. Then, the maximum likelihood classifier assigns to the leg the class label c^* of the generative model $\boldsymbol{\Psi}^c$ which maximises the likelihood of the peak locations $\mathbf{R}$,

$$c^* = \arg\max_c \log p(\mathbf{R}|\boldsymbol{\Psi}^c; \boldsymbol{\Theta}). \tag{5.32}$$

As discussed in section 5.2.2, the likelihood $p(\mathbf{R}|\boldsymbol{\Psi}^c; \boldsymbol{\Theta})$ can only be expressed as a marginal-

isation over the hidden assignment variables $\mathbf{\Omega}$,

$$\begin{aligned} c^* &= \arg\max_c \log \sum_{\mathbf{\Omega}} p(\mathbf{R}, \mathbf{\Omega} | \mathbf{\Psi}^c; \mathbf{\Theta}) \\ &= \arg\max_c \log \sum_{\mathbf{\Omega}} p(\mathbf{R} | \mathbf{\Omega}, \mathbf{X}^c, \boldsymbol{\sigma}^c; \mathbf{\Theta}) p(\mathbf{\Omega} | \mathbf{V}^c; \mathbf{\Theta}) \\ &= \arg\max_c \log E\left[p(\mathbf{R} | \mathbf{\Omega}, \mathbf{X}^c, \boldsymbol{\sigma}^c; \mathbf{\Theta}) \right]_{p(\mathbf{\Omega} | \mathbf{V}^c; \mathbf{\Theta})}, \end{aligned} \tag{5.33}$$

where E denotes expectation value.

Note that in equation (5.33) the prior distribution $p(\mathbf{\Omega}|\mathbf{V}^c;\mathbf{\Theta})$ over the assignments is now conditioned explicitly on the aspect angles $\mathbf{\Theta}$ in the test set. This is necessary because in general, the profiles in the test set will be obtained at different aspect angles than the profiles in the training set from which the generative models have been estimated. However, the elements of the visibility matrix $\mathbf{V}^c$ are the probability of visibility for the different scatterers at the aspect angles present in the training set.

The elements of $\mathbf{V}^c$ will therefore be linearly interpolated (as a function of aspect angle) to estimate a new visibility matrix $\bar{\mathbf{V}}^c$, with elements $\bar{v}^c_{nm}$ denoting the estimated probability of scatterer m being visible in the nth profile of the test set.

The prior distribution $p(\mathbf{\Omega}|\mathbf{V}^c;\mathbf{\Theta})$ is defined as

$$\begin{aligned} p(\mathbf{\Omega} | \mathbf{V}^c; \mathbf{\Theta}) &= p(\mathbf{\Omega} | \bar{\mathbf{V}}^c) \\ &= \prod_{n=1}^{N} \prod_{m=1}^{M} (\bar{v}_{nm})^{\epsilon_{nm}} (1 - \bar{v}_{nm})^{1-\epsilon_{nm}}, \end{aligned} \tag{5.34}$$

with ϵ_{nm} defined in (5.14).

Unfortunately, explicitly calculating the expectation value (5.33) is intractable due to the large number of possible assignments. A common solution is to estimate the expectation value (5.33) by averaging $p(\mathbf{R}|\mathbf{\Omega}, \mathbf{X}^c, \boldsymbol{\sigma}^c; \mathbf{\Theta})$ over a set of S samples $\mathbf{\Omega}_s$ from $p(\mathbf{\Omega}|\bar{\mathbf{V}}^c)$,

$$E\left[p(\mathbf{R} | \mathbf{\Omega}, \mathbf{X}^c, \boldsymbol{\sigma}^c; \mathbf{\Theta}) \right]_{p(\mathbf{\Omega} | \bar{\mathbf{V}}^c)} \approx \frac{1}{S} \sum_{s=1}^{S} p(\mathbf{R} | \mathbf{\Omega}_s, \mathbf{X}^c, \boldsymbol{\sigma}^c; \mathbf{\Theta}). \tag{5.35}$$

However, the distribution $p(\mathbf{\Omega}|\bar{\mathbf{V}}^c)$ is not very informative, since it does not include the information present in the peak locations $\mathbf{R}$ and the other model parameters. Therefore, a large number of samples would be required to reliably estimate the expectation value (5.35).

Fortunately, equation (5.35) can be rewritten in such a way that it can be approximated as the average over a set of samples drawn from the posterior distribution $p(\mathbf{\Omega}_s|\mathbf{R}, \mathbf{\Psi}^c; \mathbf{\Theta})$. The posterior distribution is much more informative than the prior $p(\mathbf{\Omega}|\bar{\mathbf{V}}^c)$, and so less samples are required to reliably estimate the log likelihood.

Equation (5.35) is rewritten as follows. In appendix A, equation (A.4), it is shown that the log likelihood can be written as [55]

$$\begin{aligned} \log E\left[p(\mathbf{R} | \mathbf{\Omega}, \mathbf{X}^c, \boldsymbol{\sigma}^c; \mathbf{\Theta}) \right]_{p(\mathbf{\Omega} | \bar{\mathbf{V}}^c)} = & E\left[\log p(\mathbf{R}, \mathbf{\Omega} | \mathbf{\Psi}^c; \mathbf{\Theta}) \right]_{p(\mathbf{\Omega} | \mathbf{R}, \mathbf{\Psi}^c; \mathbf{\Theta})} \\ & + H\left[p(\mathbf{\Omega} | \mathbf{R}, \mathbf{\Psi}^c; \mathbf{\Theta}) \right], \end{aligned} \tag{5.36}$$

where $H[p(\mathbf{\Omega}|\mathbf{R},\mathbf{\Psi}^c;\mathbf{\Theta})]$ denotes the entropy of $p(\mathbf{\Omega}|\mathbf{R},\mathbf{\Psi}^c;\mathbf{\Theta})$, defined by

$$\begin{aligned} H[p(\mathbf{\Omega}|\mathbf{R},\mathbf{\Psi}^c;\mathbf{\Theta})] &= -E\left[\log p(\mathbf{\Omega}|\mathbf{R},\mathbf{\Psi}^c;\mathbf{\Theta})\right]_{p(\mathbf{\Omega}|\mathbf{R},\mathbf{\Psi}^c;\mathbf{\Theta})} \\ &= -\sum_{\mathbf{\Omega}} p(\mathbf{\Omega}|\mathbf{R},\mathbf{\Psi}^c;\mathbf{\Theta}) \log p(\mathbf{\Omega}|\mathbf{R},\mathbf{\Psi}^c;\mathbf{\Theta}). \end{aligned} \tag{5.37}$$

Both the expectation value and the entropy term on the right-hand side of equation (5.36) can be approximated as an average over samples $\mathbf{\Omega}_s$, but now from the posterior distribution $p(\mathbf{\Omega}_s|\mathbf{R},\mathbf{\Psi}^c;\mathbf{\Theta})$:

$$E[\log p(\mathbf{R},\mathbf{\Omega}|\mathbf{\Psi}^c;\mathbf{\Theta})]_{p(\mathbf{\Omega}|\mathbf{R},\mathbf{\Psi}^c;\mathbf{\Theta})} \approx \frac{1}{S}\sum_s \log p(\mathbf{R},\mathbf{\Omega}_s|\mathbf{\Psi}^c;\mathbf{\Theta}), \tag{5.38}$$

$$H[p(\mathbf{\Omega}|\mathbf{R},\mathbf{\Psi}^c;\mathbf{\Theta})] \approx -\frac{1}{S}\sum_s \log p(\mathbf{\Omega}_s|\mathbf{R},\mathbf{\Psi}^c;\mathbf{\Theta}). \tag{5.39}$$

The samples used to calculate (5.39) can be obtained using the exact same sampling algorithm used during the training of the generative models, discussed in section 5.2.

5.4 Experiments

This section describes the experiments performed to test both the quality of estimated point scatter models, and the use of these models to construct a classifier as described in section 5.3.

5.4.1 Data Description

The data used in the experiments described in this section consist of simulated profiles from the five commercial aircraft described in section 2.7. The profiles were divided into a training set and a test set. The training set contained, for each aircraft, 240 profiles collected on a rectangular grid of aspect angles shown in figure 5.2. Aspect azimuth covered a range from 70 to 99 degrees, while aspect elevation ranged from 87.5 to 105 degrees.

Since commercial aircraft tend to be longer and wider than they are high, changes in aspect azimuth generally have a greater effect on range profiles than changes in aspect elevation. Therefore aspect azimuth was sampled in steps of 1 degree, while aspect elevating was sampled in steps of 2.5 degrees.

For the classification experiments an independent test set containing simulated profiles was obtained. The test set consists of five legs containing 50 range profiles each, one leg per aircraft. The profiles in leg were generated along continuous paths in aspect angle, also shown in figure 5.2.

5.4.2 Model Estimation

In the first experiment, the training set was used to estimate the parameters of a generative point scatter model for each aircraft, using the procedure described in section 5.2.

The slant ranges r_{nk} of all local maxima were estimated by fitting a quadratic function to the amplitudes of the five range bins surrounding each local maximum.

Selecting the model order, i.e. the number of scatterers to use in the model, is a difficult problem, for which no robust and efficient automated solution exists. Instead, the number of scatterers in each model was chosen a priori to be 25.

A good initialisation for the model parameters $\mathbf{X}$, $\boldsymbol{\sigma}$ and $\mathbf{V}$ is important to prevent the algorithm to converge to a local minimum. After experimenting with different initialisation

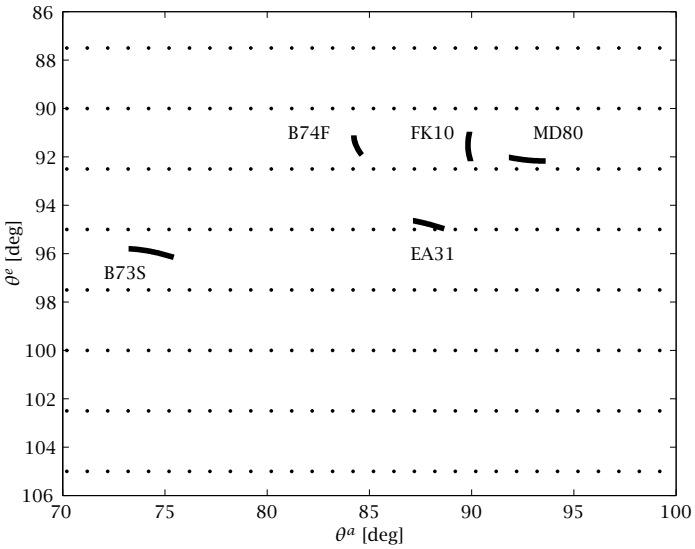


Figure 5.2: Overview of aspect angles used for generating the training and test data. The dots, arranged in a rectangular grid, denote the location of the aspect angles used in constructing the training set. Also shown are the aspect angles of the sequences in the test set, together with the corresponding aircraft code-names.

schemes, a cross-shaped formation was chosen to initialise the scatter positions. In this initialisation scheme, one-third of the scatterers are located evenly spaced across the center of aircraft's hull. The remaining scatterers were evenly spaced along a line roughly following the wings. The exact locations of these lines was determined by manual examination of the aircraft models.

The algorithm is rather insensitive to the initial values of $\boldsymbol{\sigma}$ and $\mathbf{V}$. All variances were initialised to $\boldsymbol{\sigma}_m^2 = 10\,\mathrm{m}^2$. The elements v_{nm} of $\mathbf{v}$ were initialised to 0.75, but any reasonable value seemed to work well.

The remainder of the free parameters were set as follows. The EM algorithm was run until either the scatter positions had converged, or the maximum number of 100 iterations was reached. For each profile, 10^5 assignments were sampled each iteration.

5.4.3 Classification

To test the use of the estimated models for classification, as well as the approximation of the log likelihood described in section 5.3, the following experiments were performed.

From each profile in the test set the locations of significant peaks were extracted. Then, each sequence in the test set was classified using the classifier described in section 5.3.

5.5 Results

This section discusses the results of the experiments described in the previous section. Section 5.5.1 describes the results of the model estimation experiment. Section 5.5.2 contains a more in-depth discussion of the results of estimating the model parameters for the Boeing 747-500. Finally, in section 5.4.3 the results for the classification experiment will be presented.

5.5.1 Model Estimation Results

The estimated point scatter locations and model variances are shown in figures 5.3, 5.4 and 5.5. In general, the estimated scatter locations are located on or near the body of the original aircraft model. From the figures it can be observed that scatterers located on the starboard side of the aircraft are usually in better correspondence with the model than those located on the port side.

This effect is caused by occlusion. Since all profiles in the training set were measured from approximately broadside, scatterers located on the port side of the aircraft are occluded by the aircraft body for aspect elevations close to zero. For these scatterers, less peak location samples are available, which decreases the accuracy of the location estimates.

This effect is most noticeable in the decrease in accuracy of the z-coordinate estimate of scatterers located on the port side.

5.5.2 Example: Boeing 747-500

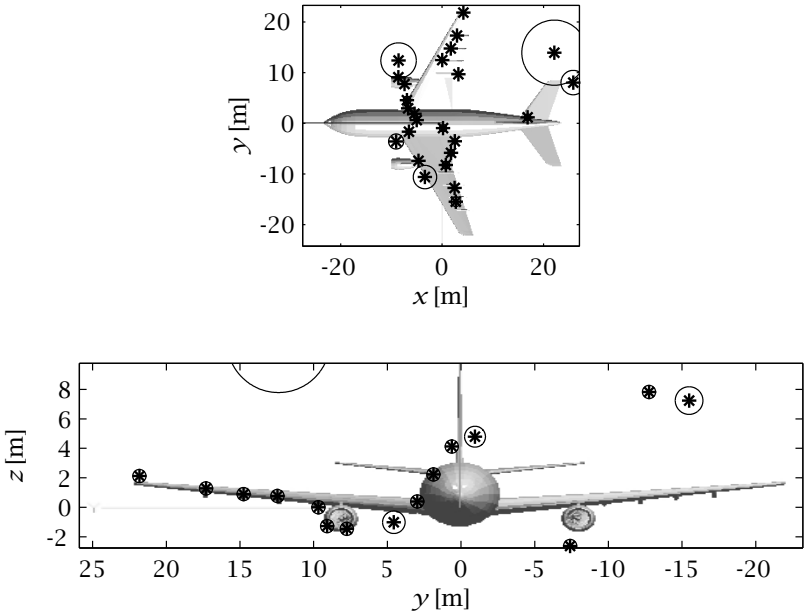
Figure 5.4(a) shows the estimated scatter positions and model variances for the Boeing 747-400, while figure 5.6 shows the corresponding visibility matrix. Figure 5.7(a) shows the peak locations extracted from the Boeing 747-400 training set, while figure 5.7(b) shows the predicted peak locations based on the estimated model parameters. The peak locations were predicted at exactly the same aspect angles as in the training set, so the two plots in figure 5.7 can be compared directly. Note that the order in which the profiles are shown in figure 5.7 is arbitrary, and has been chosen such that the data set is easily visualised. The estimation procedure is independent of any ordering.

First, note the two scatterers located on both wingtips at locations (14.7, 31.8, -2.16) and (14.9, -31.6, -3.94). They are shown in figure 5.4(a) as A and B. The locations are of these scatterers are in good, but not exact, correspondence with the actual model.The reason for the inaccuracy can be seen in figure 5.7(a). The two 'tracks' corresponding to these scatterers are easily observed at the left and right-hand side of the figure. As can be seen from this figure, peaks from other scatterers are located close to parts of the track caused by scatterers A and B. This causes uncertainty in the expectation step, which results in a less accurate estimation of the scatter positions.

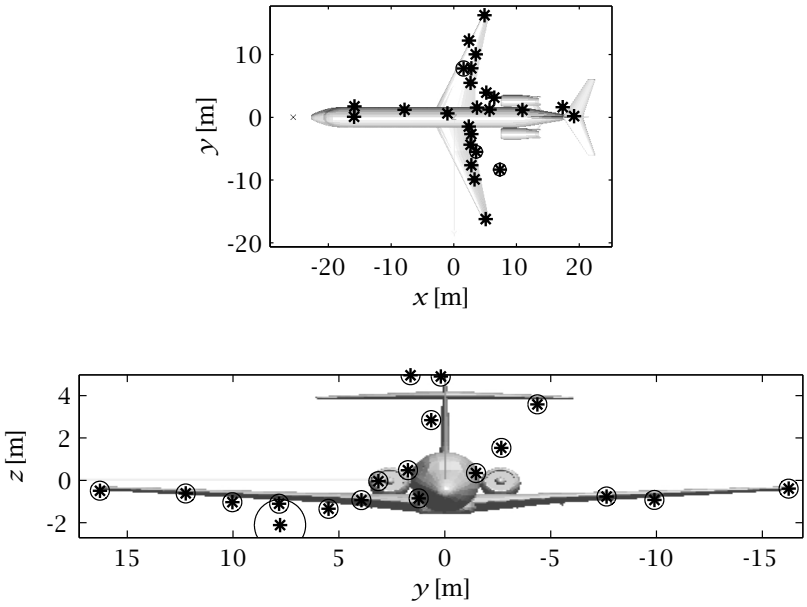
It is apparent from figure 5.7(a) that the scatterer on the left-hand wingtip is occluded in a large number of profiles. The occlusion is caused by the fact that this wingtip is occluded by the hull of the aircraft when observing the aircraft at aspect elevations close to zero. This behaviour is accurately reflect in the estimated model, as can be seen in figure 5.6. The visibility matrix for scatterer B is shown in figure 5.6 in row four, column four. The visibility matrix clearly shows scatterer B to be occluded between zero and 5 degrees elevation. Only for aspect azimuths larger than 95 degrees is it visible at 5 degrees elevation for – this occurs because the hull is not as high at the rear as it is near the middle of the aircraft. Similar occlusion effects also occur for the scatterers located around the farthest jet engine on the left hand wing, shown in columns one and two of the third row in figure 5.6.

Closer to the hull it is much harder to extract point scatterers from the peak location shown in figure 5.7(a). Especially in the region $-10 < r < 10$, many peaks occur close to each other. This makes it very difficult to establish definitive assignments in the expectation step of the algorithm. This in turn causes the point scatter location estimate to be rather inaccurate. This can easily be seen by observing the scatterers located closely to the hull in figure 5.4(a).

Although the exact location of these scatterers might not correspond perfectly to the original model, the predicted peak locations are accurate. This can be seen in figure 5.7(b), which shows the predicted peak locations at the same aspect angles as in the training set. Clear correspondences between figures 5.7(a) and 5.7(b) can be seen.

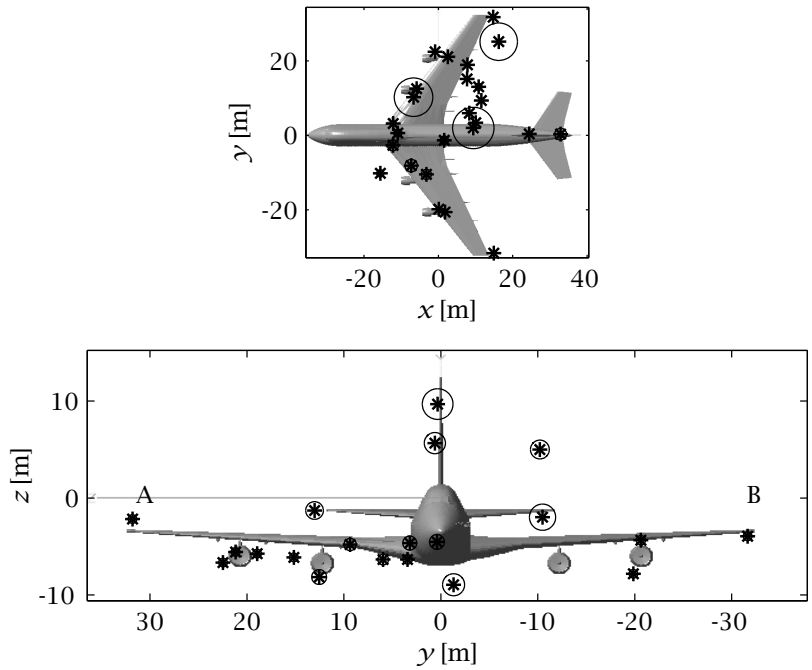


(a) Airbus 310

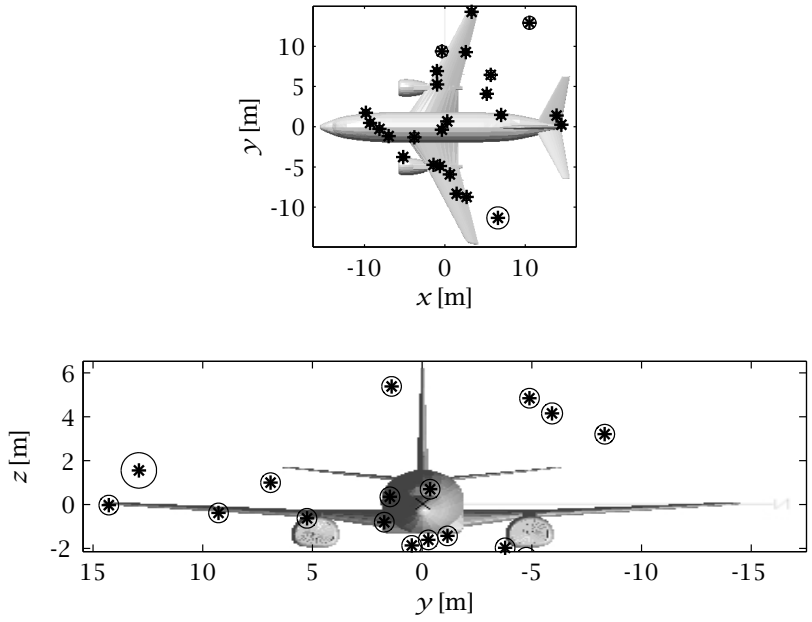


(b) McDonnell-Douglas 88

Figure 5.3: Estimated point scatter locations and model variances for the Airbus 310 and McDonnel-Douglas 88. The top figures show the projection on the xy-plane, the bottom figures the projection on the yz-plane. The stars denote the estimated scatter positions $\mathbf{x}_m$. The circles have a radius of σ_m.

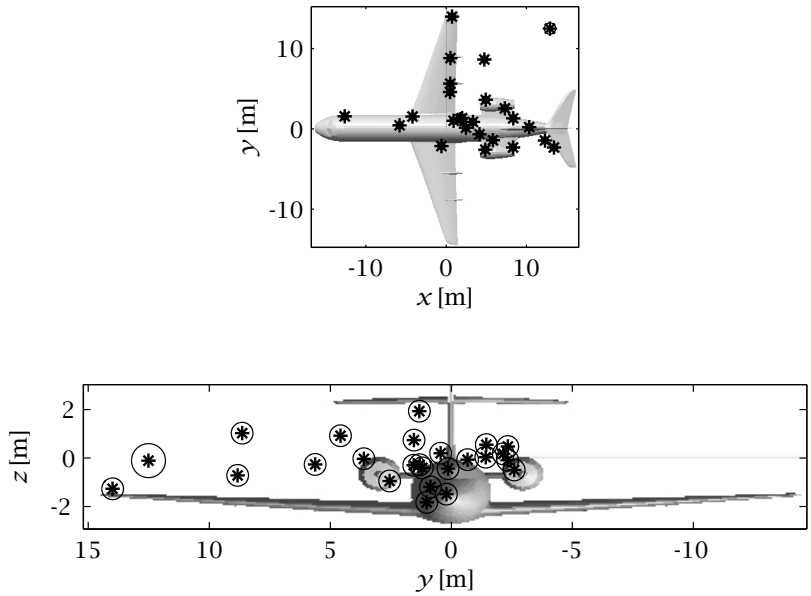


(a) Boeing 747-400



(b) Boeing 737-500

Figure 5.4: Estimated point scatter locations and model variances for the Boeing 747-400 and the Boeing 737-500. The top figures show the projection on the xy-plane, the bottom figures the projection on the yz-plane. The stars denote the estimated scatter positions $\mathbf{x}_m$. The circles have a radius of σ_m.



(a) Fokker 100

Figure 5.5: Estimated point scatter locations and model variances for the Fokker 100. The top figures show the projection on the xy-plane, the bottom figures the projection on the yz-plane. The stars denote the estimated scatter positions $\mathbf{x}_m$. The circles have a radius of σ_m.

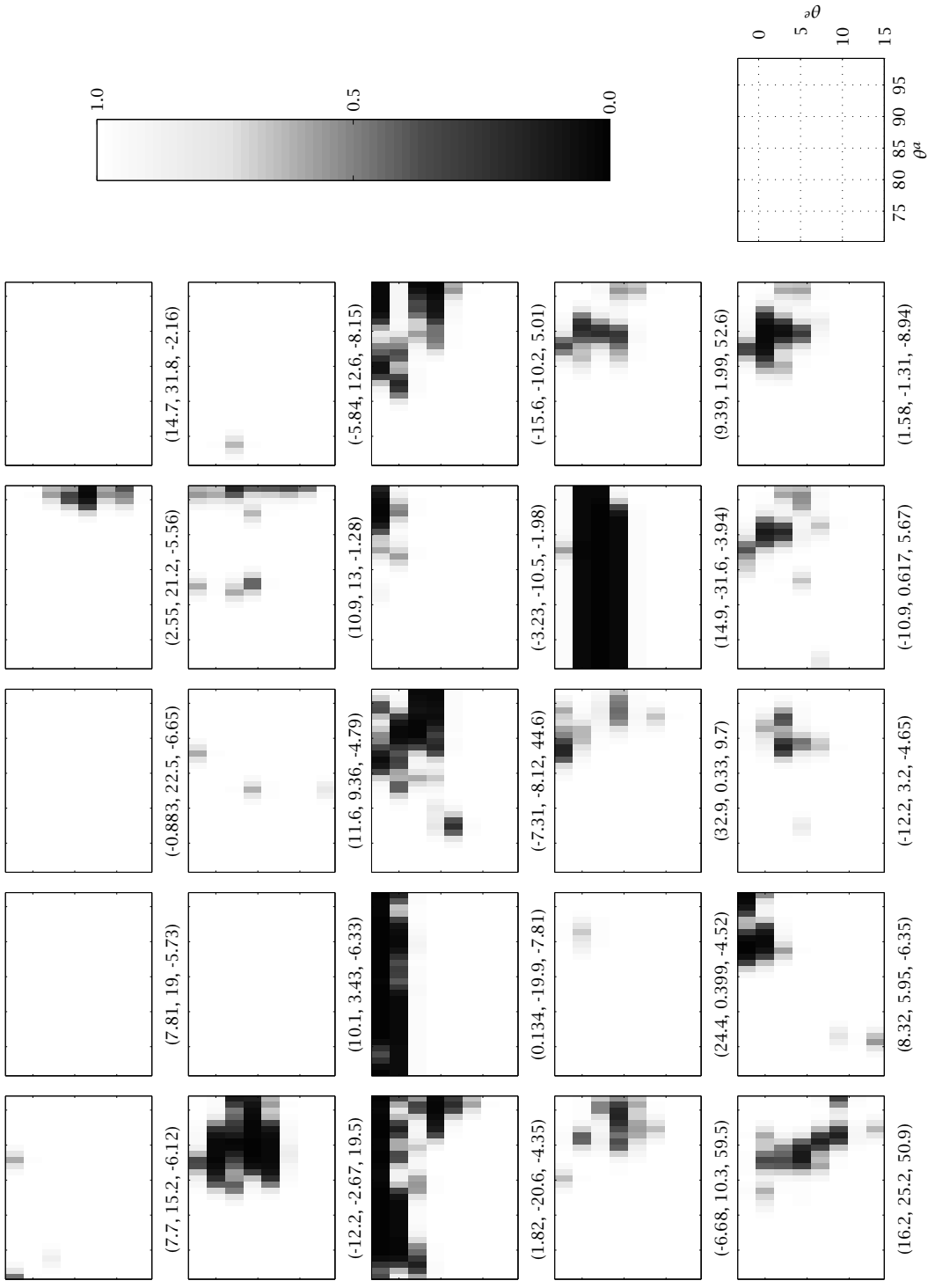
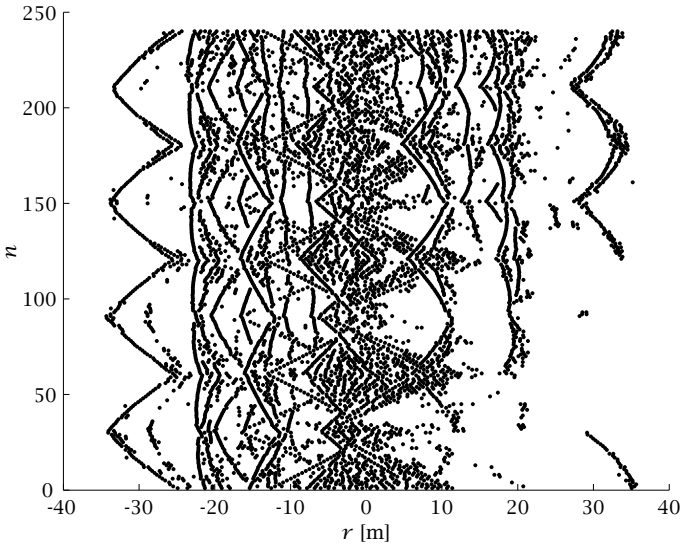
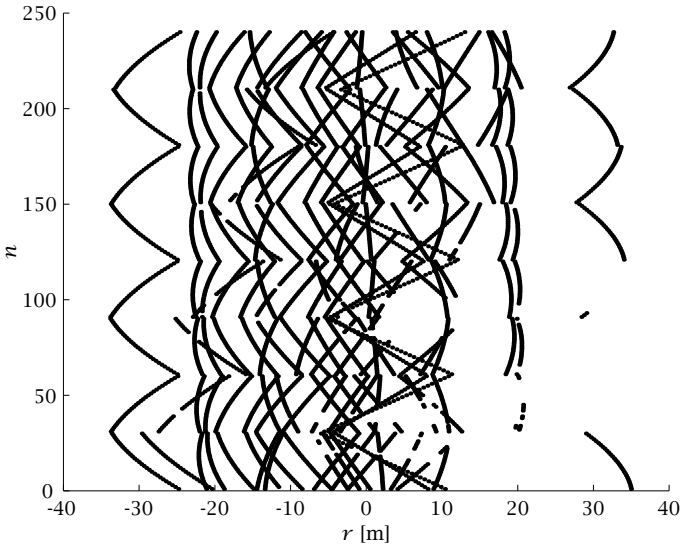


Figure 5.6: The estimated visibility matrix for the B74F. Each figure represents the estimated visibility of a scatterer in the model. The axes, shown on the lower right-hand side of this figure, are the same for each scatterer. Below each figure, the position of the corresponding scatterer is given.



(a) Peak locations in the training set of the Boeing 747-400. Some clear tracks are easily observed.



(b) Peak locations predicted by the model fitted on the above data set, at the same aspect angles as in the training set. Only peaks with an estimated probability of visibility larger than 0.5 are shown.

Figure 5.7: Observed peak locations and model predictions for the Boeing 747-400.

Class	B73S	B74F	EA31	FK10	MD80
B73S	**1.00**	0.00	0.00	0.00	0.00
B74F	0.00	**1.00**	0.00	0.00	0.00
EA31	0.00	0.00	**1.00**	0.00	0.00
FK10	0.00	0.00	0.00	**1.00**	0.00
MD80	0.00	0.00	0.00	0.00	**1.00**

Table 5.1: Classification performance of the maximum likelihood classifier for simulated range profiles. The table shows the fraction of range profiles assigned to the classes shown in the column labels. The row labels denote the true class label.

5.5.3 Classification Results

The maximum likelihood classifier assigned the correct label to each of the legs in the test set, i.e. the classification error was 0%. For completeness, the resulting confusion matrix is shown in table 5.1. The row labels in the confusion matrix show the true class of the leg. The elements of the confusion matrix are the fraction of legs assigned to the class shown in the column labels.

5.6 Conclusions

This chapter introduced a generative model of peak locations in range profiles, which is capable to generate the distribution of peak locations for a given aircraft type at any desired aspect angle.

The parameters of the generative models consist of M 3D point scatter locations $\mathbf{X}$, a set of M variances σ_m^2, and a visibility matrix V which models visibility effects caused by occlusion and speckle. It has been shown that these model parameters can be reliably estimated using an Expectation-Maximisation algorithm. Although the point scatter assumption is simple, it does fit the scatterers in the simulated training data set very well.

The parameter estimation algorithm is based on an algorithm previously described in [52], but has been improved on in several ways. In [52], the feature extraction process was performed a priori by a human operator, while in this chapter an automatic feature extraction algorithm is used. Also, in [52] the problem of occlusion is handled by selecting an a priori probability of occlusion. In this chapter, the probability of visibility of each scatterer in the model is explicitly modeled as a function of aspect angle. Furthermore, the parameters of the visibility model are estimated from the training set.

Model parameters have been estimated for each of the five aircraft shown in figure 2.4. It has been shown that the estimated model parameters show good agreement with the aircraft models. In particular, many of the estimated scatter positions coincide with clearly identifiable features on the aircraft body. Furthermore, the estimated visibility matrices clearly identify the regions in aspect angle in which scatterers are occluded.

This chapter also introduced a maximum likelihood classifier for radar range profiles, based on the previously estimated generative models. The classifier assigns a class label to the slant range positions of amplitude peaks extracted from each profile in a test set. The classifier calculates an approximation of the log likelihood of the set of peak positions for each of the estimated point scatter models. It then assigns to the test set the class label of the model for which the resulting log likelihood was the highest. Although calculation of the true log likelihood is intractable, the approximations described in section 5.3 allow us to calculate an estimate of the true log likelihood.

Since the classification used a test set containing simulated range profiles, the perfect classification results merely provide a proof of principle: model-based range profile classification from peak locations is a feasible approach. The final goal is of course to classify *measured* range profiles, using the generative models estimated from simulated range profiles. This is the topic of the next chapter.

6 Model-Based Classification of Measured Range Profiles

6.1 Introduction

In the previous chapter it was assumed that the common features between simulated and measured range profiles are the locations of peaks in the range profiles. The main goal of the current chapter is to validate that assumption by analysing the classification performance of a model-based classifier which uses only peak locations to assign class label to one or more measured range profiles.

The classifier for measured range profiles used in this chapter is very similar to the classifier for simulated range profiles discussed in chapter 5. It assigns to a set of extracted peak locations the class label of the generative point scatter model which maximises the log likelihood of the peak locations. The classifier uses the same point scatter models (estimated from simulated range profiles) as in the previous chapter.

However, classification of measured range profiles does require some modifications of the classifier described in the previous chapter, due to the differences between measured and simulated profiles. First, simulated range profiles do not suffer from measurement noise. Therefore, each peak occurring in simulated range profiles is *significant*, i.e. its presence is due to scattering from the aircraft. Measured range profiles do suffer from measurement noise (discussed in section 2.4), and so a peak in a measured range profile could be caused by either a real scattering process or by measurement noise.

Another difference between simulated and measured profiles is that for simulated profiles the exact aspect angles are known, while for measured profiles only an estimate of the aspect angle is available (see also section 2.7).

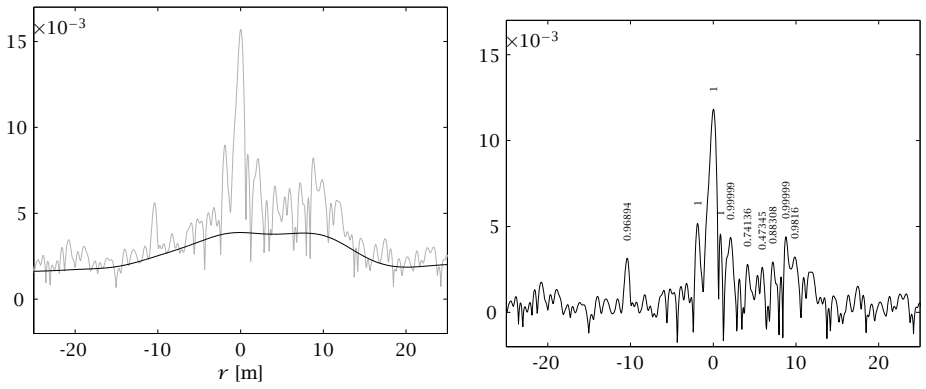
Finally, measured range profiles are subject to translational range migration (discussed in section 2.4), while simulated range profiles are not. Therefore, measured range profiles need to be aligned with respect to the point scatter models before they can be classified.

The features extracted from measured range profiles consist of a set of peak locations (as in the previous chapter), as well as a set of peak *significance values*. The significance value of a peak, defined in section 6.2, is a measure of the likelihood that a given peak is caused by a scattering process or by the measurement noise.

A generative model for the peak locations and a corresponding likelihood function for the extracted features is derived in section 6.3. It is similar to the generative model discussed in the previous chapter, but differs in two respects. First, it includes the extracted significance values as an external parameter. Second, it is modified to include the uncertainty in the aspect angle estimates and the effects of translational range migration.

Unfortunately, evaluating the derived likelihood function is intractable, and so section 6.3 also discusses a number of approximations to the earlier derived likelihood function which allow the likelihood function to be evaluated.

Section 6.4 describes three classifiers for measured range profiles. The first is a maximum likelihood classifier, which uses the generative models to classify a leg of measured range profiles. The second classifier uses a less formal similarity measure between the generative model and the extracted features. The main benefit of this classifier is that it requires less computations to classify a leg than the maximum likelihood classifier. Finally, the third classifier is a nearest neighbour classifier, which is not model-based, but directly compares measured range profiles with a training set containing simulated profiles. It's purpose in this chapter is to pro-



(a) A measured profile after the power transform.

(b) After subtracting the base amplitude.

Figure 6.1: Illustration of the base-line estimation procedure and the resulting peak significance values. Figure 6.1(a) shows a profile in grey, and the estimated base amplitude in black. Figure 6.1(b) shows the same profile after subtracting the base amplitude, as well as the significance values of a number of detected peaks.

vide a base-line classification performance to compare the performance of the other classifiers with.

Section 6.5 contains a description of the experiments performed to test the performance of the three classifiers. Section 6.6 contains the results of these experiments. Finally, section 6.7 contains the conclusions drawn from the experimental results.

6.2 Feature Extraction from Measured Profiles

The task of a classifier for measured range profiles is to assign a class label c to a leg of N range profiles $\mathbf{Y} = \{\mathbf{y}_n\}$, with aspect angle estimates $\hat{\mathbf{\Theta}} = \{\hat{\boldsymbol{\theta}}_n\}$. The profiles in the leg are assumed to be internally aligned (for instance, using maximum correlation).

The features to be extracted from the leg are the slant range locations $R = \{r_{nk}\}$ of all peaks in each profile in the leg, and corresponding significance values $\mathbf{\Gamma} = \{\gamma_{nk}\}$, where $k = 1, \cdots, K_n$, and K_n is the total number of peaks in the nth profile. A peak significance value $\gamma_{nk} = [0, 1]$ close to zero signifies that a peak is likely to be caused by the noise process, while a significance value close to 1 signifies that the peak is likely to be caused by some scattering process.

Measurement noise causes an amplitude offset in radar range profiles. This so-called *base amplitude* affects the relative amplitudes of peaks in the profile, which will be used later on to determine peak significance values. To correct for this, the base amplitude in each profile needs to be estimated, and subsequently subtracted from the profile.

The local *minima* in a range profile provide a lower bound on the base amplitude in the range bins containing the minima. To estimate the base amplitude in other range bins, the local minima are linearly interpolated, and the resulting sequence is smoothed using a low-pass Gaussian filter. An example of the resulting estimate of the base amplitude is shown in figure 6.1(a). Figure 6.1(b) shows the same profile after the the base amplitude has been subtracted.

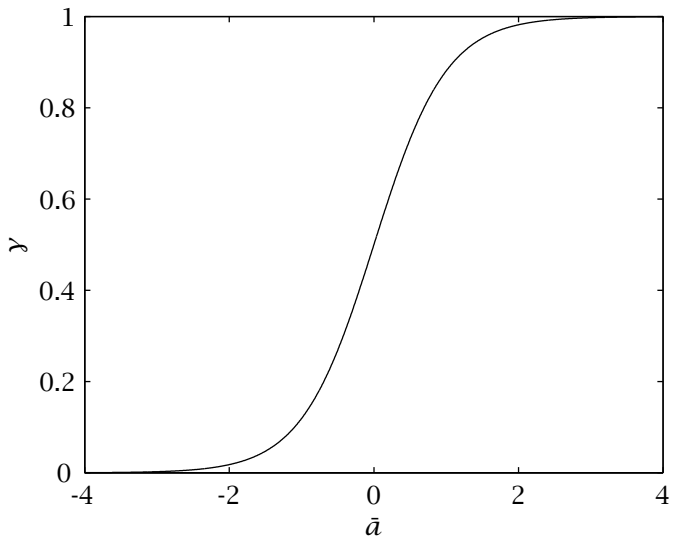


Figure 6.2: Peak significance value γ as a function of normalised peak amplitude $\bar{a}$. In this figure, $w = 2$ and $b = 3$.

After the base amplitude has been subtracted from the profile, the locations r_{nk} and amplitudes a_{nk} of all local maxima are determined by fitting a quadratic function to the amplitudes of the five range bins surrounding each local maximum.

The next step in the feature extraction process is estimating the distribution of the amplitudes of peaks caused by the noise process.

The longest aircraft used in this thesis (the Boeing 747) is about 70 metres long, while the unambiguous range interval R_u is about 107 metres. Therefore, a section of at least 30 metres in each range profile does not contain radar returns from the aircraft, i.e., it only contains noise.

By locating this region, the amplitude distribution of peaks caused by noise can be estimated. A simple heuristic to locate this region is to define it as that consecutive region of 20 metres length with the lowest average energy. The amplitude distribution of peaks caused by the noise process is then characterised by the average μ_n and standard deviation σ_n of all amplitude peaks within this region.

Finally, peak significance values $\gamma_{nk} = [0, 1]$ are defined as

$$\gamma_{nk} = \frac{1}{1 + \exp(-\bar{a}_{nk})}, \tag{6.1}$$

where

$$\bar{a}_{nk} = w \times \left(\frac{a_{nk} - \mu_n}{\sigma_n} - b \right). \tag{6.2}$$

The free parameters w and b in (6.2) determine the slope and offset of the sigmoid function (see also figure 6.2).

6.3 Derivation of Maximum Likelihood Function

6.3.1 Hidden Variables

The maximum likelihood classifier for measured range profiles assigns to a given leg $\mathbf{Y}$ the class label c^* of the generative model $\boldsymbol{\Psi}^c$ which maximises the log likelihood of the observed peak locations,

$$c^* = \arg\max_c \log p(\mathbf{R}|\boldsymbol{\Psi}^c; \hat{\boldsymbol{\Theta}}, \boldsymbol{\Gamma}). \tag{6.3}$$

Note that both the aspect angle estimates $\hat{\boldsymbol{\Theta}}$ and the peak significance values $\boldsymbol{\Gamma}$ are treated as external parameters in (6.3).

A leg of measured range profiles suffers from translational range migration, as discussed in section 2.4. In chapter 3 it was shown that profiles within a leg can be accurately aligned using maximum correlation, such that the origin of the aircraft fixed coordinate system is stationary within the leg.

The likelihood of the observed peak locations given the generative model can only be evaluated if the peaks are aligned with the model, i.e., if the origins of the two aircraft-fixed coordinate systems are at the same slant range location. The global translation in slant range a required to align the two coordinate systems is however unknown *a priori*, and so a is the second hidden variable of this problem.

The possible values of the shift a are bounded by the unambiguous range interval L (i.e. the length of the profile). The marginalisation over a of the peak likelihood is therefore given by

$$\begin{aligned} \log p(\mathbf{R}|\boldsymbol{\Psi}^c; \hat{\boldsymbol{\Theta}}, \boldsymbol{\Gamma}) &= \log \int_{a=0}^{L} p(\mathbf{R}, a|\boldsymbol{\Psi}^c; \boldsymbol{\Gamma}, \hat{\boldsymbol{\Theta}})\, da \\ &\equiv \log \int_a p(\mathbf{R}|a, \boldsymbol{\Psi}^c; \boldsymbol{\Gamma}, \hat{\boldsymbol{\Theta}}) p(a)\, da, \end{aligned} \tag{6.4}$$

where the prior distribution $p(a)$ over a is assumed to be independent of the other parameters, and is given by

$$p(a) = \frac{1}{L}. \tag{6.5}$$

Let $\mathbf{R}'$ denote the measured peak positions $\mathbf{R}$ cyclically shifted over a distance a,

$$\mathbf{R}' \equiv \{r_{nk} + a\}. \tag{6.6}$$

Equation (6.4) can then be written as

$$\log p(\mathbf{R}|\boldsymbol{\Psi}^c; \hat{\boldsymbol{\Theta}}, \boldsymbol{\Gamma}) = \log \int_a p(\mathbf{R}'|\boldsymbol{\Psi}^c; \boldsymbol{\Gamma}, \hat{\boldsymbol{\Theta}}) p(a)\, da. \tag{6.7}$$

As in section 5.2.2, the likelihood of observing the peak positions $\mathbf{R}$ is dependent on the way scatterers in the model are *assigned* to the peaks, i.e, to evaluate $p(\mathbf{R}'|\boldsymbol{\Psi}^c; \boldsymbol{\Gamma}, \hat{\boldsymbol{\Theta}})$ in (6.7) we need to know which of the observed peaks r'_{nk} are measurements of the slant range positions of which of the scatterers $\mathbf{x}_m$. The assignments are again expressed as a set of (hidden) assignment variables $\boldsymbol{\Omega}$, with the same definition as provided in section 5.2.2.

The marginalisation over $\mathbf{\Omega}$ is again given by the sum over all possible values of $\mathbf{\Omega}$,

$$\begin{aligned}\log p(\mathbf{R}|\mathbf{\Psi}^c;\hat{\mathbf{\Theta}},\mathbf{\Gamma}) &= \log \int_a \sum_{\mathbf{\Omega}} p(\mathbf{R}',\mathbf{\Omega}|\mathbf{\Psi}^c;\mathbf{\Gamma},\hat{\mathbf{\Theta}})p(a)\,da \\ &= \log \int_a \sum_{\mathbf{\Omega}} p(\mathbf{R}'|\mathbf{\Omega},\mathbf{\Psi}^c;\mathbf{\Gamma},\hat{\mathbf{\Theta}})p(\mathbf{\Omega},a|\mathbf{\Psi}^c;\mathbf{\Gamma},\hat{\mathbf{\Theta}})p(a)\,da.\end{aligned} \tag{6.8}$$

6.3.2 Definition of Probability Distributions

Now that the hidden variables have been included in the likelihood function, the probability distributions in (6.8) can be defined. The probability distributions are similar to the distributions defined in section 5.2.4, but are adapted to include the uncertainty in the aspect angle estimates $\hat{\mathbf{\Theta}}$ and the peak significance values $\mathbf{\Gamma}$.

The distribution $p(\mathbf{R}'|\mathbf{\Omega},\mathbf{\Psi}^c;\hat{\mathbf{\Theta}})$ can be factorised, similar to equation 5.15, as

$$p(\mathbf{R}'|\mathbf{\Omega},\mathbf{\Psi}^c;\hat{\mathbf{\Theta}}) = \prod_{n=1}^{N}\prod_{k=1}^{K_n} p(r'_{nk}|\omega_{nk},\mathbf{\Psi}^c;\hat{\boldsymbol{\theta}}_n). \tag{6.9}$$

If the kth scatterer in the nth profile has been assigned to the mth scatterer, $\omega_{nk} = m$, the peak position distribution $p(r'_{nk}|\omega_{nk} = m,\mathbf{\Psi}^c;\hat{\boldsymbol{\theta}}_n)$ is assumed to be Gaussian.

In the previous chapter, the mean of this Gaussian distribution was defined as the slant range $r(\boldsymbol{\theta},\mathbf{x}_m)$ of the mth scatterer, defined in equation (2.5) as

$$r(\boldsymbol{\theta},\mathbf{x}_m) = \mathbf{s}(\boldsymbol{\theta})\cdot\mathbf{x}_m. \tag{6.10}$$

The variance of this distribution was the model parameter σ_m^2.

However, the uncertainty in the aspect angle estimate $\hat{\boldsymbol{\theta}}$ induces an uncertainty in the slant range $r(\hat{\boldsymbol{\theta}},\mathbf{x}_m)$. Assuming a zero-mean Gaussian error distribution on $\hat{\boldsymbol{\theta}}$, with covariance matrix $\mathbf{\Sigma}_{\boldsymbol{\theta}}$, the resulting variance $\sigma_{\boldsymbol{\theta}}$ on $r(\hat{\boldsymbol{\theta}},\mathbf{x}_m)$ is given by

$$\sigma_{\hat{\boldsymbol{\theta}}}^2 = (\mathbf{r}_\perp)^{\mathrm{T}}\,\Sigma_{\boldsymbol{\theta}}\,\mathbf{r}_\perp, \tag{6.11}$$

$$\mathbf{r}_\perp = (r_\perp^a, r_\perp^e)^{\mathrm{T}}, \tag{6.12}$$

where the cross ranges $r_\perp^a$ and $r_\perp^e$ are defined in equation (2.7).

Therefore, the mean of the Gaussian distribution $p(r'_{nk}|\omega_{nk} = m,\mathbf{\Psi}^c;\hat{\boldsymbol{\theta}}_n)$ is itself Gaussian distributed, with mean $r(\hat{\boldsymbol{\theta}},\mathbf{x}_m)$ and variance $\sigma_{\hat{\boldsymbol{\theta}}}^2$. The resulting peak location distribution is therefore given by

$$p(r'_{nk}|\omega_{nk} = m,\mathbf{X}^c;\hat{\mathbf{\Theta}}) = \mathcal{N}\left(r'_{nk};r(\hat{\boldsymbol{\theta}}_n,\mathbf{x}_m),\sigma_m^2+\sigma_{\hat{\boldsymbol{\theta}}}^2\right), \tag{6.13}$$

i.e., the variance of the peak location distribution is the sum of the model variance σ_m^2 and the variance $\sigma_{\hat{\theta}}^2$ induced by the variance of the aspect angle estimate $\hat{\boldsymbol{\theta}}$.

If the kth peak has *not* been assigned to any scatterers, $\omega_{nk} = 0$, the peak location likelihood is given by

$$p(r'_{nk}|\omega_{nk} = 0,\mathbf{X}^c;\hat{\mathbf{\Theta}}) = \frac{1}{L}, \tag{6.14}$$

which is identical to (5.12).

In section 5.2.4, equation (5.15), the prior distribution $p(\mathbf{\Omega}|\mathbf{\Psi}^c;\mathbf{\Theta})$ was defined as

$$\begin{aligned} p(\mathbf{\Omega}|\mathbf{\Psi}^c;\mathbf{\Theta}) &= p(\mathbf{\Omega}|\mathbf{V}^c) \\ &= \prod_{n=1}^{N} p(\boldsymbol{\omega}_n|\mathbf{v}_n^c) \\ &= \prod_{n=1}^{N}\prod_{m=1}^{M} (v_{nm}^c)^{\epsilon_{nm}}(1-v_{nm}^c)^{1-\epsilon_{nm}}, \end{aligned} \tag{6.15}$$

with $\mathbf{V}^c$ the visibility matrix of class c. Each element v_{nm}^c of the visibility matrix $\mathbf{V}^c$ denotes the probability of scatterer m being present in the nth profile of the simulated training set, where the profile was obtained at a specific aspect angle $\boldsymbol{\theta}_n$.

In general, the measured range profiles in the test set will be measured at different aspect angles $\hat{\boldsymbol{\theta}}_n$. To obtain an estimate $\bar{\mathbf{V}}$ of the probability of visibility at the estimated aspect angles $\hat{\boldsymbol{\theta}}_n$, the elements of V will be linearly interpolated as a function of aspect angle, as described in section 5.3.

The definition (6.15) will now be adapted to include the significance values $\mathbf{\Gamma}$ as an additional parameter.

The prior probability of assigning scatterer m with visibility $\bar{v}_{nm}$ to peak in profile n at location r'_{nk} with peak significance γ_{nk} is given by

$$p(\omega_{nk} = m|\bar{v}_{nm}^c, \gamma_{nk}) = \bar{v}_{nm}\gamma_{nk}. \tag{6.16}$$

The prior probability of *not* assigning the mth scatterer to any peak, $w_{nk} = 0$, in the nth profile is given by

$$p(\omega_{nk} = 0|\bar{v}_{nm}^c, \gamma_{nk}) = (1-\bar{v}_{nm}^c), \tag{6.17}$$

which is the same definition as used in chapter 5.

Let κ_{nk} be a binary variable indicating whether the kth peak has been assigned to *any* scatterer in the nth profile,

$$\kappa_{nk} = \begin{cases} 1 & \text{if } \omega_{nk} \neq 0 \\ 0 & \text{if } \omega_{nk} = 0 \end{cases}. \tag{6.18}$$

With this definition, the total prior probability of a set of assignments $\mathbf{\Omega}$, given a visibility matrix $\mathbf{V}^c$ and peak significance values $\mathbf{\Gamma}$, can be written as

$$\begin{aligned} p(\mathbf{\Omega}|\mathbf{V}^c, \mathbf{\Gamma}; \hat{\mathbf{\Theta}}) &= \prod_{n=1}^{N} p(\boldsymbol{\omega}_n|\bar{\mathbf{v}}_n^c, \boldsymbol{\gamma}_n) \\ &= \prod_{n=1}^{N}\left(\prod_{m=1}^{M}(\bar{v}_{nm}^c)^{\epsilon_{nm}}(1-\bar{v}_{nm}^c)^{1-\epsilon_{nm}}\prod_{k=1}^{K_n}(\gamma_{nk})^{\kappa_{nk}}\right). \end{aligned} \tag{6.19}$$

Combining the assumptions and definitions presented in this section, the peak position log likelihood can be written as a nested expectation value,

$$\begin{aligned} \log p(\mathbf{R}|\mathbf{\Psi}^c; \hat{\mathbf{\Theta}}, \mathbf{\Gamma}) &= \log \int_a \sum_{\mathbf{\Omega}} p(\mathbf{R}'|\mathbf{\Omega}, \mathbf{\Psi}^c; \hat{\mathbf{\Theta}}) p(\mathbf{\Omega}|\mathbf{V}^c, \mathbf{\Gamma}; \hat{\mathbf{\Theta}}) p(a)\, da \\ &= \log E\left[E\left[p(\mathbf{R}'|\mathbf{\Omega}, \mathbf{\Psi}^c; \hat{\mathbf{\Theta}})\right]_{p(\mathbf{\Omega}|\mathbf{V}^c, \mathbf{\Gamma}; \hat{\mathbf{\Theta}})}\right]_{p(a)}. \end{aligned} \tag{6.20}$$

6.3.3 Evaluation of Expected Log Likelihood

Unfortunately, the expected log likelihood (6.20) can not be easily evaluated. First, the expectation value over $p(a)$ involves an integral over a which can not be performed analytically. Also, the expectation value over $p(\mathbf{\Omega}|\mathbf{V}^c, \mathbf{\Gamma}; \hat{\mathbf{\Theta}})$ involves a sum over an exponential number of possible assignments, which is computationally very expensive. This section therefore presents some approximations which allow (6.20) to be evaluated within a reasonable time.

A common approach to calculating expectation values is to sample from the distribution over which the expectation value is taken, and approximate the expectation value as the average posterior likelihood taken over the set of samples.

Unfortunately, this approach is not applicable to the expectation value over $p(a)$ in (6.20) for two reasons. The first reason is that the prior $p(a) = 1/L$ is flat, and therefore not informative. As a consequence, a large number of samples would be required to reliable approximate the expectation value. The second reason is that, for each sample from $p(a)$, the marginalisation over $\mathbf{\Omega}$ would have to be performed, which would take a prohibitive amount of time.

Instead, equation (6.20) is approximated by estimating an optimal value of the global shift a_c^*, and replacing the original definition $p(a) = 1/L$ with

$$p(a) = \delta(a - a_c^*). \tag{6.21}$$

Equation (6.20) then reduces to

$$\begin{aligned} \log p(\mathbf{R}|\mathbf{\Psi}^c; \hat{\mathbf{\Theta}}, \mathbf{\Gamma}) &= \log \int_a \sum_{\mathbf{\Omega}} p(\mathbf{R}'|\mathbf{\Omega}, \mathbf{\Psi}^c; \hat{\mathbf{\Theta}}) p(\mathbf{\Omega}|\mathbf{V}^c, \mathbf{\Gamma}; \hat{\mathbf{\Theta}}) \delta(a - a_c^*)\, da \\ &= \log \sum_{\mathbf{\Omega}} p(\mathbf{R}^c|\mathbf{\Omega}, \mathbf{\Psi}^c; \hat{\mathbf{\Theta}}) p(\mathbf{\Omega}|\mathbf{V}^c, \mathbf{\Gamma}; \hat{\mathbf{\Theta}}) \end{aligned} \tag{6.22}$$

where $R^c = \{r_{nk} + a_c^*\}$.

The optimal shift a_c^* for a given class c is estimated using a *observed peak signature matrix* $\mathbf{P}$, constructed from the extracted features, and *model peak signature matrices* $\mathbf{Q}^c$, constructed from the generative model of each class c. A peak signature matrix is an $N \times D$ matrix, where N is the number of profiles in the leg, and D the number of range bins in each profile. The rows $\mathbf{p}_n$ of $\mathbf{P}$ are called *observed peak signatures*; the rows $\mathbf{q}_n^c$ of $\mathbf{Q}^c$ are called *model peak signatures*.

Observed peak signatures are constructed from the extracted features as follows. First, a function G is defined as the weighted sum of K_n Gaussian kernels, with means r_{nk} and constant variances σ_k^2 (which is a free parameter). The weighting factors for each Gaussian kernel are the peak significance values γ_{nk}:

$$G_n(r) = \sum_k \gamma_{nk} \mathcal{N}(r_{nk}, \sigma_k^2). \tag{6.23}$$

Then, the observed peak signature $\mathbf{p}_n$ is constructed by sampling G at the slant ranges r_d of the original range bins, and normalising the resulting vector:

$$p_{nd} = \frac{G_n(r_d)}{\sum_d G_n r_d}. \tag{6.24}$$

Model peak signatures are constructed similarly. In this case, the function G^c is a weighted sum of M Gaussian kernels, with means $r(\hat{\boldsymbol{\theta}}, \mathbf{x}_m^c)$ and variances $\sigma_m^2 + \sigma_\theta^2$. The weighting factors in this case are the scatter visibilities v_{nm}^c:

$$G_n^c(r) = \sum_m v_{nm}^c \mathcal{N}(r(\hat{\boldsymbol{\theta}}, \mathbf{x}_m^c), \sigma_m^2 + \sigma_\theta^2). \tag{6.25}$$

As before, the model peak signature $\mathbf{q}_n^c$ is constructed by sampling G^c at the slant ranges r_d of the range bins in the measured range profile, and normalising the resulting vector:

$$q_{nd} = \frac{G_n(r_d)}{\sum_d G_n r_d}. \tag{6.26}$$

Figure 6.3 provides an overview of the construction of peak signatures.

From each pair of peak signature matrices $\mathbf{P}$ and $\mathbf{Q}^c$, an optimal shift in slant range a_c^* is estimated by cyclically shifting the rows of $\mathbf{P}$, and calculating the resulting 'inner product' χ of $\mathbf{P}$ and $\mathbf{Q}^c$,

$$\begin{aligned} a_c^* &= \arg\max_a \mathbf{P}^a \cdot \mathbf{Q}^c \\ &\equiv \arg\max_a \sum_n \mathbf{p}_n^a \cdot \mathbf{q}_n. \end{aligned} \tag{6.27}$$

After the extracted peak locations have been shifted over the optimal shift a_c^*, the remaining expectation value over $p(\mathbf{\Omega}|\mathbf{V}^c, \mathbf{\Gamma})$ is given by

$$\log p(\mathbf{R}|\mathbf{\Psi}^c; \hat{\mathbf{\Theta}}, \mathbf{\Gamma}) = \log E\left[p(\mathbf{R}^c|\mathbf{\Omega}, \mathbf{\Psi}^c; \hat{\mathbf{\Theta}})\right]_{p(\mathbf{\Omega}|\mathbf{V}^c, \mathbf{\Gamma}; \hat{\mathbf{\Theta}})}. \tag{6.28}$$

Equation (6.28) has the exact same form as equation (5.36) – consequently, it can be evaluated using the same procedure as discussed in section 5.3.

6.4 Classifiers for Measured Range Profiles

This section discusses three different classifiers for measured range profiles. Two of the classifiers are model-based. The third is a nearest neighbour classifier which operates on range profiles directly.

In describing the different classifiers it is assumed that the data to be classified, i.e., the *test set*, consists of a leg $\mathbf{Y} = \{\mathbf{y}_n\}$ containing N profiles, and a set of corresponding aspect angle estimates $\hat{\mathbf{\Theta}} = \{\hat{\boldsymbol{\theta}}_n\}$.

6.4.1 Maximum Likelihood Classifier

This section summarises the feature extraction process and subsequent calculation of the peak location likelihood described in the previous sections.

The maximum likelihood classifier for measured range profiles assigns to the profiles in the test set the class label c^* of the model $\mathbf{\Psi}^c$ which maximises the log likelihood of the observed peak locations. The maximum likelihood classifier can operate in three 'modes': 1-look classification, majority voting, and full leg classification.

In 1-look classification mode, the classifier assigns a class label c_n^* to each profile $\mathbf{y}_n$ in $\mathbf{Y}$ separately, according to

$$c_n^* \arg\max_c \log p(\mathbf{r}_n|\mathbf{\Psi}^c; \hat{\boldsymbol{\theta}}_n, \boldsymbol{y}_n). \tag{6.29}$$

In majority voting mode, the classifier assigns a single class label c^* to all profiles in the leg. In this mode, c^* is the class label which the majority of profiles have been assigned in the 1-look classification mode. In the case of a tie, the class assigns the leg to an unknown class, labeled NOMA (for No Majority).

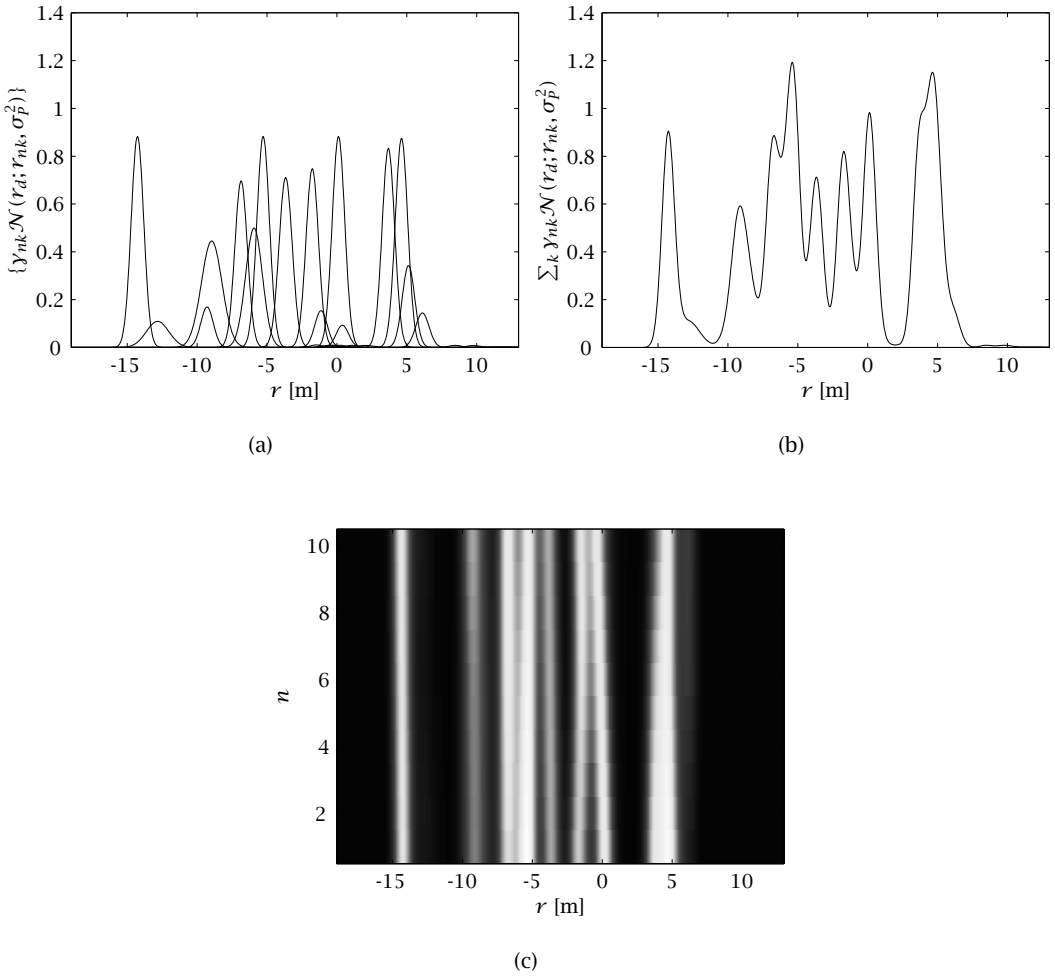


Figure 6.3: Construction of a peak location signature. The predicted peak positions, variances and visibility probabilities (shown in figure 6.3(a)) are combined to construct a peak location signature (shown in figure 6.3(b)). The resulting peak signature matrix is shown in figure 6.3(c).

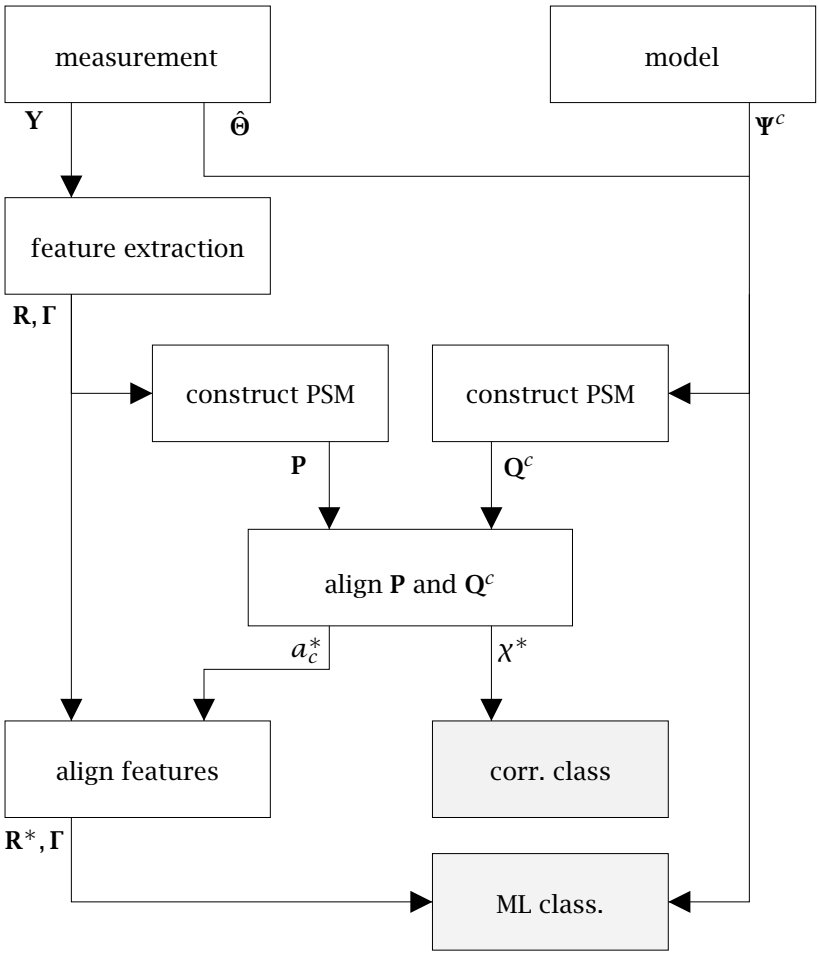


Figure 6.4: Overview of the processing steps of the maximum likelihood classifier and the maximum correlation classifier.

Finally, in full leg classification mode, the classifier assigns a single class label c^* to all profiles in the leg, based on the total likelihood of extracted peak locations,

$$c^* = \arg\max_c \log p(\mathbf{R}|\mathbf{\Psi}^c; \hat{\mathbf{\Theta}}, \mathbf{\Gamma}). \tag{6.30}$$

A summary of the processing involved in maximum likelihood classification in full leg classification mode is presented below. (The extension to the other modes of operation is straightforward.) A schematic overview is shown in figure 6.4.

1. Features $\mathbf{R}$ and $\mathbf{\Gamma}$ are extracted from the leg. From these features a peak signature matrix, $\mathbf{P}$, is constructed.

2. By combining the aspect angle estimates $\hat{\mathbf{\Theta}}$ with each point scatter model, a peak signature matrix $\mathbf{Q}^c$ is constructed for each class c.

3. For each class c, the optimal shift a_c^* of the extracted peak locations $\mathbf{R}$ is estimated by searching for the shift which maximises the inner product of $\mathbf{P}$ and $\mathbf{Q}^c$.

4. A set of assignment samples is drawn from $p(\mathbf{\Omega}|\mathbf{R}^c, \mathbf{\Psi}^c; \hat{\mathbf{\Theta}}, \mathbf{\Gamma})$, where R^c is the set of extracted peak locations cyclically shifted by a_c^*.

5. From the set of sampled assignments, the log likelihood for each class is estimated using (6.28).

6. The test set is assigned the class label c^* of the model which produces the maximum log likelihood.

6.4.2 Heuristic Alternative: Maximum Correlation

A drawback of the maximum likelihood classifier is that the computational requirements are rather high, due to the necessity of sampling the assignment distribution. This section introduces an alternative model-based classifier for measured range profiles with more modest computational requirements.

Section 6.3.3 describes the peak signature matrices $\mathbf{P}$ and $\mathbf{Q}^c$, which are used to calculate an optimal global shift a_c^* of the peak locations for each class c. However, the maximum correlation between $\mathbf{P}$ and $\mathbf{Q}^c$ is itself a measure of similarity between the extracted features and the generative model – the more peaks they have in common, the higher the maximum correlation will be in general.

A simple and quick alternative to maximum likelihood classification is therefore to use the maximum value of the cross-correlation between $\mathbf{P}$ and $\mathbf{Q}^c$ as a measure of similarity between the extracted features and the generative model.

The maximum correlation classifier can operate in the same three modes as the maximum likelihood classifier: 1-look classification, majority voting, and full leg classification.

In 1-look classification mode, the classifier assigns a class label c_n^* to each profile $\mathbf{y}_n$ in $\mathbf{Y}$ separately, according to

$$c_n^* \arg\max_c \max_a \mathbf{p}_n^a \cdot \mathbf{q}_n^c. \tag{6.31}$$

In majority voting mode, the classifier assigns a single class label c^* to all profiles in the leg. In this mode, c^* is the class label which the majority of profiles have been assigned in the 1-look classification mode. In the case of a tie, the class assigns the leg to an unknown class, labeled NOMA (for No Majority).

Finally, in full leg classification mode, the classifier assigns a single class label c^* to all profiles in the leg, based on the total maximum correlation between $\mathbf{P}$ and $\mathbf{Q}^c$,

$$c^* = \arg\max_c \max_a \mathbf{P}^a \cdot \mathbf{Q}^c. \tag{6.32}$$

Although this classifier lacks a solid statistical foundation, it is significantly faster then the maximum likelihood classifier, since the time-consuming process of sampling the possible assignments is no longer required. Furthermore, the computations necessary to use the maximum correlation classifier are in fact identical to the first set of computations necessary to use the maximum likelihood classifier (see also figure 6.4). Therefore, the maximum correlation classifier can be used to obtain a preliminary classification while the maximum likelihood classifier is still being computed.

6.4.3 Nearest Neighbour Classifier

The nearest neighbour classifier is commonly used in many classification tasks. It is used in this chapter to provide a base-line classification performance, with which the performance of the other two classifier are compared.

The nearest neighbour classifier is not model-based, but instead compares the test profiles $\mathbf{y}_n$ directly with the 'training set' of simulated profiles.

Let $\mathbf{Z} = \{\mathbf{z}_t\}$, $t = 1, \ldots, T$ denote the simulated profiles in the training set, and let c_t denote the class label of $\mathbf{z}_t$. Finally, let $\boldsymbol{\theta}_n$ denote the aspect angle of $\mathbf{z}_n$.

To increase the reliability of the nearest neighbour classifier, a test profile $\mathbf{y}_n$ should only be compared to training profiles $\mathbf{z}_t$ whose aspect angles $\boldsymbol{\theta}_t$ are within the expected error of the aspect angle estimate $\hat{\boldsymbol{\theta}}_n$ of $\mathbf{y}_n$. Since the expected error on aspect angle estimates is about 5 degrees (see section 2.7), a test profile $\mathbf{y}_n$ is only compared to training profiles $\mathbf{z}_{t'}$ for which the distance between $\hat{\boldsymbol{\theta}}_n$ and $\boldsymbol{\theta}_n$ is below 5 degrees,

$$t' = \{t \in 1, \ldots, T \;\Big|\; |\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_t| < 5\}. \tag{6.33}$$

The nearest neighbour classifier can operate in two modes: 1-look classification and majority voting.

In 1-look classification mode, the classifier assigns a class label c_n^* to each profile $\mathbf{y}_n$ in $\mathbf{Y}$ separately, according to

$$c_n^* = c_{t^*}, \tag{6.34}$$

$$t^* = \arg\min_{t'} \text{SED}(\mathbf{y}_n, \mathbf{z}_{t'}). \tag{6.35}$$

In majority voting mode, the classifier assigns a single class label c^* to all profiles in the leg. In this mode, c^* is the class label which the majority of profiles have been assigned in the 1-look classification mode. In the case of a tie, the class assigns the leg to an unknown class, labeled NOMA (for No Majority).

6.5 Experiments

This section describes the classification experiments performed to test the three classifiers for measured range profiles discussed in the previous section.

6.5.1 Data Description

The data set contained simulated and measured range profiles from the five different commercial aircraft used throughout this thesis (shown in figure 2.4).

The radar returns were Hamming weighted and zero-padded to 2048 elements. Finally, a Box-Cox transformation, defined in equation (2.27), was applied to each available profile. The free parameter η was chosen to be $\eta = 0.2$.

The simulated profiles were used to estimate a generative point scatter model for each aircraft as described in the previous chapter – the same models were used in the experiments described in section 5.4.

The measured range profiles were measured in the ORFEO measurement campaign, described in section 2.7. Not all available range profiles were used: only range profiles measured at aspect azimuth θ^a ranging from 70 to 105 degrees and aspect elevation θ^e ranging from 87.5 to 102 degrees were used.

The (rather long) legs in the original ORFEO data were divided into sub-legs containing 10 profiles each. Table 6.1 shows the total number of available legs for each aircraft. Figure 6.5 shows the aspect angles of each leg. The profiles in each sub-leg were then aligned with respect to each other using maximum correlation, as described in chapter 3.

Aircraft code	Number of legs
B73S	102
B74F	17
EA31	34
FK10	22
MD80	87

Table 6.1: Number of available legs for each aircraft.

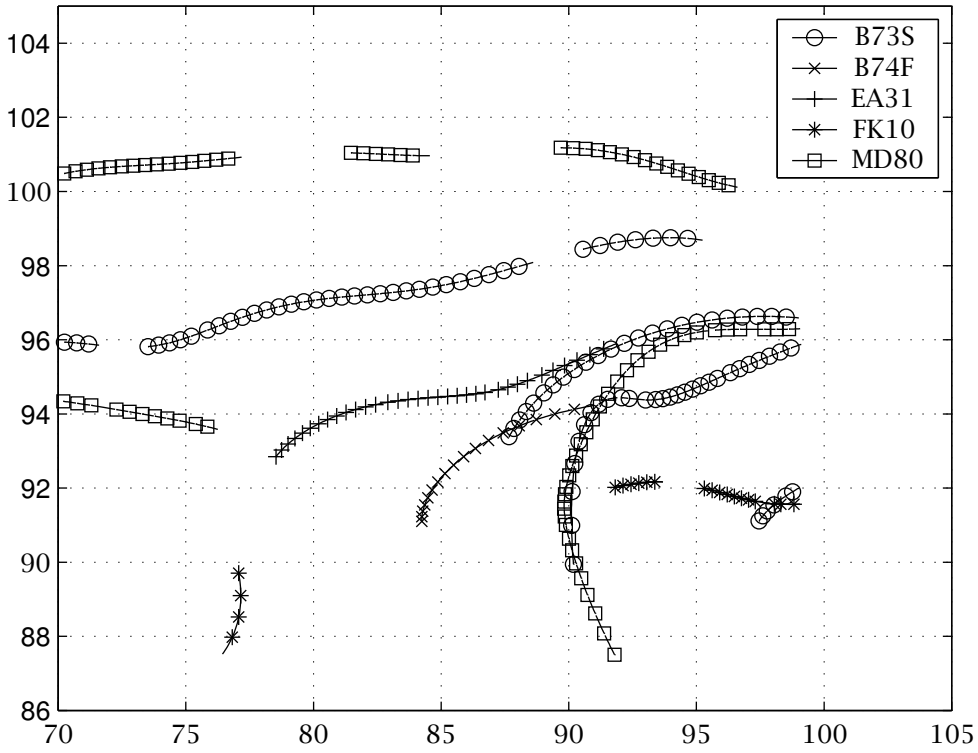


Figure 6.5: Aspect angles of available measured data.

Finally, the peak locations $\mathbf{R}$ and peak significance values $\mathbf{\Gamma}$ were extracted from the profiles in each leg, according to the feature extraction described in section 6.2. The variance of the Gaussian filter used in the baseline estimation procedure was $\sigma^2 = 20\ \mathrm{m}^2$. The free parameters w and b defined in (6.2) were chosen to be $w = 2$ and $b = 3$. Note that these parameters are not estimated from the data, but are instead chosen *a priori*.

6.5.2 Classification

Each classifier (in all its operating modes) discussed in section 6.4 was used to classify the test set.

6.6 Results and Discussion

The results of each classification experiment are presented in tables 6.2, 6.3 and 6.4. The results are presented as confusion matrices. The row labels denote the true class of the legs, while the column labels denote the class assigned by the classifier. The elements of the confusion matrix represent the fraction of the available data assigned to the class denoted in the column label.

Table 6.2 contains the confusion matrices for each operating mode of the maximum likelihood classifier. The full leg classification mode of the maximum likelihood classifier clearly outperforms both other modes, both in terms of overall classification error as well as in terms of the classification error for four of the five classes – only the classification error for the FK10 is slightly lower in the majority voting mode.

In all modes, the B74F is almost always classified correctly. This is not surprising, since the B74F is by far the largest aircraft in the data set.

The results for the MD80 vary the most over the different operating modes of the maximum likelihood classifier. Although the classification error is the same for both the 1-look classification mode and the majority voting mode, the relatively large number of legs which are classified as NOMA shows that the peak likelihood fluctuates rapidly within the legs of the MD80. The full leg classification mode assigns the correct class label to almost all legs previously classified as NOMA.

Table 6.3 contains the confusion matrices for each operating mode of the maximum correlation classifier. Overall, the performance of the maximum correlation classifier is comparable, mode by mode, to the performance of the maximum likelihood classifier. As with the maximum likelihood classifier, the 1-look classification mode is results in the poorest performance, and the full leg classification mode results in the best performance of the three modes.

In fact, the results of the full leg classification mode of the maximum correlation for each aircraft separately are slightly better than for the same mode of the maximum likelihood classifier – the overall error is however most affected by the large number of legs of the B73S in the test set. Since the computational requirements for the maximum correlation classifier are so much lower than for the maximum likelihood classifier, one could argue that the maximum correlation classifier is the more attractive choice in practice.

Table 6.4(a) contains the confusion matrix for the two operating modes of the nearest neighbour classifier. Both the 1-look classification mode and the majority voting mode of the nearest neighbour classifier outperforms the similar modes of the two model-based classifiers, mainly due to the increase in performance of the B73S.

Since there is no full leg classification mode for the nearest neighbour classifier, one can only compare the full leg classification modes of the model-based classifiers to the majority voting mode of the nearest neighbour classifier. The overall performance in these cases is roughly the same for all three classifiers, but again, the overall performance is heavily influenced by the large number of legs of the B73S.

Class	B73S	B74F	EA31	FK10	MD80
B73S	**0.62**	0.02	0.10	0.12	0.14
B74F	0.00	**0.96**	0.04	0.01	0.00
EA31	0.01	0.41	**0.56**	0.00	0.03
FK10	0.17	0.00	0.02	**0.63**	0.18
MD80	0.23	0.01	0.11	0.08	**0.57**

(a) 1-Look Classification. Overal error: 0.38

Class	B73S	B74F	EA31	FK10	MD80	NOMA
B73S	**0.75**	0.00	0.05	0.08	0.07	0.05
B74F	0.00	**1.00**	0.00	0.00	0.00	0.00
EA31	0.00	0.38	**0.62**	0.00	0.00	0.00
FK10	0.09	0.00	0.00	**0.77**	0.14	0.00
MD80	0.16	0.00	0.09	0.05	**0.57**	0.13

(b) Majority Voting. Overal error: 0.31

Class	B73S	B74F	EA31	FK10	MD80
B73S	**0.79**	0.00	0.05	0.09	0.07
B74F	0.00	**1.00**	0.00	0.00	0.00
EA31	0.00	0.32	**0.68**	0.00	0.00
FK10	0.09	0.00	0.00	**0.73**	0.18
MD80	0.17	0.00	0.06	0.06	**0.71**

(c) Full leg. Overal error: 0.24

Table 6.2: Confusion matrices for the maximum likelihood classifier. The tables show the fraction of range profiles assigned to the classes shown in the column labels. The row labels denote the true class label.

Class	B73S	B74F	EA31	FK10	MD80
B73S	**0.60**	0.01	0.05	0.22	0.13
B74F	0.00	**0.96**	0.02	0.01	0.01
EA31	0.01	0.34	**0.58**	0.00	0.07
FK10	0.21	0.00	0.03	**0.63**	0.13
MD80	0.21	0.01	0.06	0.15	**0.57**

(a) 1-Look Classification. Overal error: 0.39

Class	B73S	B74F	EA31	FK10	MD80	NOMA
B73S	**0.71**	0.00	0.04	0.16	0.07	0.03
B74F	0.00	**1.00**	0.00	0.00	0.00	0.00
EA31	0.00	0.32	**0.59**	0.00	0.03	0.06
FK10	0.18	0.00	0.00	**0.77**	0.05	0.00
MD80	0.11	0.00	0.05	0.11	**0.68**	0.05

(b) Majority Voting. Overal error: 0.29

Class	B73S	B74F	EA31	FK10	MD80
B73S	**0.77**	0.00	0.04	0.14	0.05
B74F	0.00	**1.00**	0.00	0.00	0.00
EA31	0.00	0.24	**0.76**	0.00	0.00
FK10	0.14	0.00	0.05	**0.77**	0.05
MD80	0.15	0.00	0.06	0.06	**0.74**

(c) Full leg. Overal error: 0.26

Table 6.3: Confusion matrices for the correlation classifier. The tables show the fraction of range profiles assigned to the classes shown in the column labels. The row labels denote the true class label.

Class	B73S	B74F	EA31	FK10	MD80
B73S	**0.88**	0.00	0.00	0.09	0.03
B74F	0.00	**0.99**	0.00	0.01	0.00
EA31	0.15	0.00	**0.71**	0.04	0.10
FK10	0.16	0.00	0.00	**0.58**	0.25
MD80	0.29	0.00	0.00	0.19	**0.51**

(a) 1-Look Classification. Overal error: 0.28

Class	B73S	B74F	EA31	FK10	MD80	NOMA
B73S	**0.96**	0.00	0.00	0.03	0.01	0.00
B74F	0.00	**1.00**	0.00	0.00	0.00	0.00
EA31	0.21	0.00	**0.68**	0.03	0.09	0.00
FK10	0.14	0.00	0.00	**0.73**	0.14	0.00
MD80	0.30	0.00	0.00	0.18	**0.52**	0.00

(b) Majority Voting. Overal error: 0.24

Table 6.4: Confusion matrices for the nearest neighbour classifier. The table shows the fraction of range profiles assigned to the classes shown in the column labels. The row labels denote the true class label.

Overall it can be concluded that there is no significant difference in the performance of the three classifiers on the data set available for the experiments.

6.7 Conclusions

In the previous chapter it was shown that it was possible to estimate generative point scatter models from simulated range profiles, and that these models could be used to reliably classify simulated range profiles. Furthermore, it was assumed that the slant ranges of amplitude peaks are the common features between measured and simulated range profiles.

This chapter described two new model-based classifiers for measured range profiles: a maximum likelihood classifier and a maximum correlation classifier. Both use the same generative model defined in chapter 5 to assign a class label to a leg of measured range profiles,

The comparable performance of the model-based classifiers and the nearest neighbour classifier show that the common features between simulated and measured range profiles are indeed the slant ranges of amplitude peaks. Furthermore, these common features can be modeled from simulated data. The resulting generative models are capable of discriminating between measured profiles of different aircraft types, although not with 100% accuracy.

Of the three classifiers described in this chapter, the maximum correlation classifier provides a good compromise between theoretical rigor and practical requirements. While it lacks a formal statistical foundation, its performance is comparable to the maximum likelihood classifier, while requiring significantly less computations. Furthermore, it is model-based, and it does not require a large continuously available training set, which the nearest neighbour classifier does.

The maximum likelihood classifier is theoretically sound – it is a proper maximum likelihood classifier, based on a simple but physically relevant model of radar scattering. Therefore, future research should focus on further developing this classifier.

Instead of assigning scatterers to peaks, an interesting approach would be to assign scatterers to *tracks* of amplitude peaks in a given leg. This would dramatically reduce the number of possible assignments, which is the main reason for the large computational requirements of the maximum likelihood classifier.

Furthermore, the generative model underlying the maximum likelihood classifier could be extended to include more complex scattering mechanisms than simple point scattering.

7 Conclusions

Aircraft classification from radar range profiles is a challenging task due to the three problems identified in chapter 1: the need for translation invariant classification, the limited accuracy of available aspect angle estimates, and the difficulty of obtaining sufficient training data. These problems are addressed in the previous chapters in this thesis. This final chapter provides an overview of the main results as well as directions for possible future research.

Chapter 3 addressed the problem of translation invariant classification by introducing the Smoothed Zero Phase Representation for radar range profiles. This representations allows for fast translation invariant classification, while not discarding any potentially discriminative information from the range profiles.

The main benefits of using the Smoothed Zero Phase Representation compared to other approaches to translation invariant range profile classification are twofold. First, the method aligns range profiles in a pre-processing step, and so it becomes possible to apply statistical feature extraction methods on the data set. This is not possible with many other approaches, which obtain translation invariance by using a translation invariant similarity measure to classify range profiles. The second benefit is that translation invariant classification is faster than classification using the common approach of using a nearest neighbour classifier based on the Sliding Euclidean Distance. Classification of a leg containing M range profiles requires roughly M times less alignment-related computations than classification using the SED.

Using the ZPR or SZPR does result in a decrease in classifier performance compared to classification using the SED. However, by combining range profile alignment using the SZPR with subsequent statistical feature extraction could result in better classification performance. Whether such an approach outperforms classification using the SED would be an interesting topic for future research.

The aspect angle at which an aircraft is observed heavily influences range profile measurements. The effects of speckle, rotational range migration and occlusion cause significant range profile variability. Therefore, knowledge about the aspect angle at which a range profile is obtained can significantly increase classification accuracy. Furthermore, information on aspect angles also benefit many HRR signal processing techniques, such as ISAR imaging.

Aspect angle estimates from tracking data suffer from both systematic and random errors (see chapter 2). Chapter 4 introduced a method for increasing the accuracy of aircraft rotation estimates obtained from tracking data. On the basis of a point scatter model a filter was developed, which uses a combination of the original rotation estimates and phase information extracted from range profiles to update the rotation estimates.

To extract the necessary phase information from the range profiles an extensive feature extraction method was developed, which involves tracking peaks detected in a leg of range profiles, and correcting extracted phases for the effects of translational and rotational range migration.

Since no ground truth on aspect angles are available for the range profiles measured during the ORFEO campaign, the experiments were necessarily restricted to simulated profiles. However, the effects of measurement noise and translational range migration were artificially added to the simulated range profiles, and so the resulting data set is similar to actual HRR measurements.

It has been shown that the filter does indeed improve aircraft rotation estimates, provided that the signal-to-noise ratio is sufficiently high (larger than 0 dB). For smaller signal-to-noise

ratios, the algorithm fails to provide more accurate rotation estimates, resulting in an improvement factor of 1.

The main area of future research regarding this method is its application on actual measured range profiles, for which a ground truth on aspect angles is available.

In this thesis simulated range profiles were used to construct classifiers for measured range profiles, which requires solving two problems. Since simulated range profiles are significantly different from measured range profiles, the classifier should base its decision solely on those features of range profiles which are present in both simulated and measured range profiles. Therefore, the first problem to be solved is to identify those common features. The second problem is actually constructing a classifier capable of classifying the extracted features.

Chapter 5 assumed that the common features present in both simulated and measured range profiles are the locations (i.e. the slant ranges) of the local maxima occurring in range profiles. On the basis of that assumption, a generative model describing the distribution of peak positions was defined. The model is based on a simple point scatter model of radar scattering. It's parameters include the three dimensional positions of a set of point scatterers, as well as a visibility matrix which models the effects of speckle and occlusion.

For each aircraft in the training set model parameters were estimated using an Expectation-Maximisation algorithm. It was shown that the resulting scatter positions and visibility matrices are in good agreement with the original CAD models used to generate the simulated training set.

Finally, the performance of a maximum likelihood classifier based on this generative model was shown to result in perfect classification accuracy on a test set containing simulated range profiles. This is an important proof of principle: model-based range profile classification from peak locations is a feasible approach.

An interesting topic for future research is the use of more sophisticated model of radar scattering as the basis of a generative model for peak locations. Although the current point scattering model performs well in practice, more sophisticated models could further increase the expressive power of the generative model. Also, it would be interesting to look into the problem of model order selection, i.e. the problem of deciding the number of scatterers present in the model.

Chapter 6 used the generative models obtained in chapter 5 to construct two model-based classifiers for measured range profiles. The first classifier is a maximum-likelihood classifier very similar to the one presented in chapter 5. The second classifier is also model-based, but uses a similarity measure instead of the maximum likelihood to classify range profiles.

Measured range profiles are subjected to measurement noise and translational range migration. Chapter 6 presented an algorithm to assign a significance value to peaks extracted from measured range profiles, which is a measure of the likelihood that a peak was caused by actual scattering, or by the noise process. The generative model was modified to include these significance values in the definition of peak position likelihood. The problem of translational range migration was solved by first aligning the peaks extracted from the measurements with the distribution predicted by the model before classification. The performance of the resulting classifiers was shown to be comparable to the performance of a nearest neighbour classifier, which proves that peak locations are indeed the correct features for comparing simulated and measured range profiles.

The identification of the common features between simulated and measured range profiles is the most important result presented in this thesis, which leads to many possibilities for further research.

The accuracy of the peak detection algorithms and the assignment of significance values might be improved by exploiting the fact that usually a leg of profiles is available. In this thesis, peaks are detected on a profile by profile basis. However, one could also try to detect tracks of peaks in the leg (as was done in chapter 4). The significance value of a peak could

then be related to the likelihood of it being part of track.

Another benefit of using tracks of peaks is that this could decrease the computational requirements of the maximum likelihood classifier. Instead of assigning scatterers to peaks, one would assign scatterers to tracks instead. This would dramatically reduce the number of possible assignments, which is the main reason for the large computational requirements of the maximum likelihood classifier.

The obvious long-term goal of any research on non-cooperative aircraft recognition is actual application in radar systems. The main question which has to be answered before such an application is possible is whether sufficient classification performance can be achieved when many more aircraft types are included in the data set. The work presented in this thesis has hopefully provided the next step towards answering that question.

A Expectation-Maximisation Optimisation

This appendix provides more details on the Expectation-Maximisation algorithm described in section 5.2.3. It closely follows the description presented in [55].

First, let $q(\mathbf{\Omega})$ denote an *arbitrary* distribution over the hidden variables $\mathbf{\Omega}$. The log likelihood $\log p(\mathbf{R}|\mathbf{\Psi})$ can be written as

$$\begin{aligned}
\log p(\mathbf{R}|\mathbf{\Psi}) &= \log \sum_{\mathbf{\Omega}} p(\mathbf{R}, \mathbf{\Omega}|\mathbf{\Psi}) \\
&= \log \sum_{\mathbf{\Omega}} q(\mathbf{\Omega}) \frac{p(\mathbf{R}, \mathbf{\Omega}|\mathbf{\Psi})}{q(\mathbf{\Omega})} \\
&\geq \sum_{\mathbf{\Omega}} q(\mathbf{\Omega}) \log \frac{p(\mathbf{R}, \mathbf{\Omega}|\mathbf{\Psi})}{q(\mathbf{\Omega})} \\
&\equiv F[q, \mathbf{\Psi}],
\end{aligned} \tag{A.1}$$

where in the third step Jensen's inequality is used. The lower bound $F[q, \Psi]$ is called the *free energy*.

The Expectation-Maximisation algorithm is an iterative algorithm, which in each iteration finds an updated set of model parameters which maximise the free energy. Each iteration consists of an Expectation step and a Maximisation step.

Let t be the number of the current iteration. In the Expectation step, the model parameters $\mathbf{\Psi}^t$ are fixed. By writing the free energy as

$$\begin{aligned}
F[q, \mathbf{\Psi}^t] &= \sum_{\mathbf{\Omega}} q(\mathbf{\Omega}) \log \frac{p(\mathbf{R}, \mathbf{\Omega}|\mathbf{\Psi}^t)}{q(\mathbf{\Omega})} \\
&= \sum_{\mathbf{\Omega}} q(\mathbf{\Omega}) \log p(\mathbf{R}|\mathbf{\Psi}^t) + \sum_{\mathbf{\Omega}} q(\mathbf{\Omega}) \log \frac{p(\mathbf{\Omega}|\mathbf{R}, \mathbf{\Psi}^t)}{q(\mathbf{\Omega})} \\
&= E[\log p(\mathbf{R}|\mathbf{\Psi}^t)]_{q(\mathbf{\Omega})} - \sum_{\mathbf{\Omega}} q(\mathbf{\Omega}) \log \frac{q(\mathbf{\Omega})}{p(\mathbf{\Omega}|\mathbf{R}, \mathbf{\Psi}^t)} \\
&= \log p(\mathbf{R}|\mathbf{\Psi}^t) - D[q(\mathbf{\Omega}) || p(\mathbf{\Omega}|\mathbf{R}, \mathbf{\Psi}^t)],
\end{aligned} \tag{A.2}$$

where D denotes the Kullback-Leibler divergence. The optimal choice $q^t(\mathbf{\Omega})$ for q is given by

$$q^t(\mathbf{\Omega}) = p(\mathbf{\Omega}|\mathbf{R}, \mathbf{\Psi}^t), \tag{A.3}$$

since in that case $D = 0$ and the free energy is equal to the log likelihood.

In the Maximisation step, $q^t(\mathbf{\Omega})$ is fixed, and the free energy is maximised as a function of $\mathbf{\Psi}$. A different expansion of the free energy is given by

$$\begin{aligned}
F[q^t, \mathbf{\Psi}) &= \sum_{\mathbf{\Omega}} q^t(\mathbf{\Omega}) \log \frac{p(\mathbf{R}, \mathbf{\Omega}|\mathbf{\Psi})}{q^t(\mathbf{\Omega})} \\
&= \sum_{\mathbf{\Omega}} q^t(\mathbf{\Omega}) \log p(\mathbf{R}, \mathbf{\Omega}|\mathbf{\Psi}) - \sum_{\mathbf{\Omega}} q^t(\mathbf{\Omega}) \log q^t(\mathbf{\Omega}) \\
&= E[\log p(\mathbf{R}, \mathbf{\Omega}|\mathbf{\Psi})]_{q^t(\mathbf{\Omega})} + H[q^t(\mathbf{\Omega})],
\end{aligned} \tag{A.4}$$

where $H[q^t(\mathbf{\Omega})]$ denotes the entropy of $q^t(\mathbf{\Omega})$. Since q^t is fixed, the entropy of q^t is independent of $\mathbf{\Psi}$. Therefore, the new optimal model parameters $\mathbf{\Psi}^{t+1}$ are those that maximise the first part of (A.4), i.e., the expected log likelihood:

$$\mathbf{\Psi}^{t+1} = \arg\max_{\mathbf{\Psi}} E[\log p(\mathbf{R}, \mathbf{\Omega}|\mathbf{\Psi})]_{p(\mathbf{\Omega}|\mathbf{R},\mathbf{\Psi}^t)}. \tag{A.5}$$

Bibliography

[1] R.O. Duda and P.E. Hart. *Pattern Classification and Scene Analysis*. John Wiley & Sons, 1973.

[2] Keinosuke Fukunaga. *Introduction to Statistical Pattern Recognition*. Academic Press, 2nd edition, 1990.

[3] Andrew R. Webb. *Statistical Pattern Recognition*. Arnold, 1999. ISBN 0-340-74164-3.

[4] Anil K. Jain, Robert P.W. Duin, and Jianchang Mao. Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1):4–37, 2000.

[5] Christoph Nieuwoudt and Elizabeth C. Botha. Relative performance of correlation-based and neural-network-based classifiers of aircraft using synthetic radar range profiles. *Opt. Eng.*, 39(5):1245–1251, May 2000.

[6] Y.T. Chan, K.C. Ho, and S.K. Wong. Aircraft identification from RCS measurement using an orthogonal transform. *IEE Proc.-Radar, Sonar, Navig.*, 47(2):93–102, April 2000.

[7] R. van der Heiden, F.C.A. Groen, and L.J. van Ewijk. Aircraft recognition with radar range profiles using a synthetic database. In *RTO SET Symposium on High Resolution Radar Techniques*, 1999.

[8] Richard A. Mitchell and John J. Westerkamp. Robust statistical feature based aircraft identification. *IEEE Transactions on Aerospace and Electronic Systems*, 35(3):1077–1094, 1999.

[9] Dzung Tri Pham. Applications of unsupervised clustering algorithms to aircraft identification using high range resolution radar. In *Proceedings of the IEEE 1998 National Aerospace and Electronics Conference*, pages 228–235, 1998.

[10] Kie B. Eom. Time-varying autoregressive modeling of HRR radar signatures. *IEEE Transactions on Aerospace and Electronic Systems*, 35(3):974–988, July 1999.

[11] R. van der Heiden. *Aircraft Recognition with Radar Range Profiles*. PhD thesis, University of Amsterdam, 1998.

[12] A. Zyweck and R.E. Bogner. Radar target classification of commercial aircraft. *IEEE Transactions on Aerospace and Electronics Systems*, 32(2):598–606, april 1996.

[13] S. Hudson and D. Psaltis. Correlation filters for aircraft identification from radar range profiles. *IEEE Transactions on Aerospace and Electronics Systems*, 29(3):741–748, July 1993.

[14] Jim P. Ballard and Tom P. Leonard. A maximum likelihood range-profile classifier for maritime targets. In *RTO SCI Symposium on "Non-Cooperative Air Target Identification Using Radar"*, 1998.

[15] V. Gouaillier and L. Gagnon. Ship silhouette recognition using principal component analysis. In *SPIE Proc. 3164, Applications of Digital Image Processing XX*, 1997.

[16] Hossam Osman, Li Pan, Steven D. Blostein, and Langis Gagnon. Classification of ships in airborne SAR imagery using backpropagation neural networks. In *SPIE Proc. 3161, Radar Processing, Technology and Applications II*, pages 126–136, 1997.

[17] John S. Baras and Sheldon I. Wolk. Wavelet based progressive classification of high range resolution radar returns. In *SPIE Proceedings 3391, Wavelet Applications V*, pages 546–558, 1994.

[18] A.R. Webb. Gamma mixture models for target revognition. *Pattern Recognition*, 33(12):2045–2054, 2000.

[19] De-Shuang Huang. A novel forward-backward smoothing-based learning subspace method for recognition of radar targets. *International Journal of Pattern Recognition and Artificial Intelligence*, 13(1):65–83, 1999.

[20] Dave Gross, John Greenewald, James Schmitz, Robert Williams, and John Westerkamp. High range resolution ground moving target ATR using advanced space-based SAR/MTI concepts. In *AIAA Space Technology Conference & Exposition*, pages 1–10, September 1999.

[21] Yosuke Nakano, Yoshihisa Hara, and Jun Saito. Radar target recognition system using 3-D mathematical model. In *SPIE Conference on Automatic Target Recognition VIII*, pages 83–91, 1998.

[22] Xejun Liao and Zheng Bao. Two new categories of shift-invariant features of high-resolution radar range profiles. In *Proceedings of ICSP '98*, pages 1485–1488, 1998.

[23] K.P.H Cambien. Aircraft aspect angle estimation with radar tracking data. Master's thesis, Delft University of Technology, August 2000.

[24] Merril I. Skolnik. *Introduction to Radar Systems*. McGraw-Hill, 2nd edition, 1985.

[25] D.R. Wehner. *High-Resolution Radar*. Artech House, 2nd edition, 1994. ISBN 0-89006-727-9.

[26] J.B. Keller. Geometrical theory of diffraction. *J. Opt. Society of America*, pages 116–130, January 1962.

[27] Eugene F. Knott, John F. Shaeffer, and Michael T. Tuley. *Radar Cross Section*. Artech House, 2nd edition, 1993. ISBN 0-89006-618-3.

[28] August W. Rihaczek and Stephen J. Hershkowitz. *Theory and Practice of Radar Target Identification*. Artech House, 2000. ISBN 1-58053-081-8.

[29] Lee C. Potter and Randolph L. Moses. Attributed scattering centres for SAR ATR. *IEEE Transactions on Image Processing*, 6(1):79–91, January 1997.

[30] M.J. Gerry, L.C. Potter, I.J. Gupta, and A. van der Merwe. A parametric model for synthetic aperture radar measurements. *IEEE Transactions on Antennas and Propagation*, 47(7):1179–1188, July 1999.

[31] Lee C. Potter, Da-Ming Chiang, Rob Carrière, and Michael J. Gerry. A GTD-based parametric model for radar scattering. *IEEE Transactions on Antennas and Propagation*, 43(10):1058–1067, October 1995.

[32] P. Swerling. Probability of detection for fluctuating targets. *IRE Trans.*, IT-6:269–308, April 1960.

[33] F.E. Nathanson. *Radar Design Principles*. McGraw-Hill, 1969. Chapter 5.

[34] Fredric J. Harris. On the use of windows for harmonic analysis with the Discrete Fourier Transform. *Proc. of the IEEE*, 66(1):51–83, 1978.

[35] R.W. Hamming. *Digital Filters*. Dover Publications, 3rd edition, 1989. ISBN 0-486-65088-X.

[36] G.E.P. Box and D.R. Cox. An analysis of transformation. *Journal of the Royal Statistical Society, Series B*, 26:211–252, 1964.

[37] R.M. Sakia. The Box-Cox transformation technique: a review. *The Statistician*, 41:169–178, 1992.

[38] R. van der Heiden and F.C.A. Groen. The Box-Cox metric for nearest neighbour classification improvement. *Pattern Recognition*, 30(2), 1997.

[39] Les Piegl. On NURBS: a survey. *IEEE Computer Graphics and Applications*, 1991.

[40] D.J. Andersh, M. Hazlett, S.W. Lee, and D.D Reeves. XPATCH: A high-frequency electromagnetic-scattering prediction code and environment for complex three-dimensional objects. *IEEE Antennas and Propagation Magazine*, 36:65–69, 1994.

[41] M.G.E. Brand. Radar signature analysis and prediction by physical optics and ray-tracing. The RAPPORT code for RCS prediction. Technical Report FEL–95–A097, TNO-FEL, 1995.

[42] R. van der Heiden and J. de Vries. The ORFEO measurement campaign. Technical Report FEL-96-A073, TNO FEL, 1996.

[43] P. Kosir, R. DeWall, and R.A. Mitchell. A multiple measurement approach for feature alignment. In *Proceedings of the IEEE 1995 National Aerospace and Electronics Conference NAECON 1995*, pages 94–101, 1995. ISBN: 0-7803-2667-9.

[44] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in C++*. Cambridge University Press, 2nd edition, 2002.

[45] T. Pajdla and V. Hlaváč. Zero phase representation of panoramic images for image based localization. In F. Solina and A. Leonardis, editors, *8th International Conference on Computer Analysis of Images and Patterns*, pages 550–557. Springer Berlin, 1999. ISBN 3-540-66366-5.

[46] Hungh-Chih Chiang, Randolph L. Moses, and Lee C. Potter. Model-based Bayesian feature matching with application to synthetic aperture radar target recognition. *Pattern Recognition*, 31:1539–1553, 2001.

[47] Joseph A. O'Sullivan, Steven P. Jacobs, and Vikas Kedia. Stochastic models and performance bounds for pose estimation using high resolution radar data. In *SPIE Conference on Algorithms for Synthetic Aperture Radar Imagery*, pages 576–587, 1998.

[48] Steven P. Jacobs and Joseph A. O'Sullivan. Automatic target recognition using sequences of high resolution radar range-profiles. *IEEE Transactions on Aerospace and Electronic Systems*, 36(2):364–381, apr 2000.

[49] Guy L. Scott and Christopher Longuet-Higgins. An algorithm for associating the features of two images. *Proc. R. Soc. Lond. B*, 244:21–26, April 1991.

[50] Richard Allen Mitchell. *Robust High Range Resolution Radar Target Identification Using A Statistical Feature Based Classifier with Feature Level Fusion*. PhD thesis, School of Engineering, Univeristy of Dayton, 1997.

[51] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Statist. Soc. Ser. B*, 59(3):1–38, 1977.

[52] Frank Dellaert, Steven Seitz, Charles Thorpe, and Sebastian Thrun. Structure from motion without correspondence. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'00)*, June 2000.

[53] Peter J. Green. On use of the EM for penalized likelihood estimation. *Journal of the Royal Statistical Society. Series B (Methodological)*, 52(3):443–452, 1990.

[54] Radford M. Neal. Probabilistic inference using Markov chain Monte Carlo methods. Technical Report CRG-TR-93-1, Dept. of Computer Science, University of Toronto, September 1993.

[55] R.M. Neal and G.E. Hinton. A view of the EM algorithm that justifies incremental, sparse, and other variants. In M.I. Jordan, editor, *Learning in Graphical Models*, pages 355–368. Kluwer Academic Publishers, 1998.

[56] Markus Svensén. *GTM: The Generative Topographic Mapping*. PhD thesis, Aston University, Birmingham, UK, 1998.

[57] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller. Equations of state calculations by fast computing machine. *Journal of Chemical Physics*, 21:1087–1091, 1953.

[58] W. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57:97–109, 1970.

Summary

The ability to quickly and reliably identify aircraft is an important aspect of air traffic safety. Civilian air traffic controllers need to be constantly updated on the status of aircraft moving through the local airspace. In military scenarios, the need to reliably identify aircraft is even more stringent, since erroneous identification could easily result in friendly fire incidents.

A common technique for identification of military aircraft is Identification Friend Foe (IFF), which relies on the aircraft providing a correct response to a *challenge* transmitted by a ground station. An aircraft not capable of correctly answering the challenge is assumed to be hostile (or at least not friendly).

Unfortunately this assumption does not always hold. In recent history many incidents have shown that friendly aircraft are not always capable of providing a correct response, be it due to hardware failure or human error.

The subject of this thesis is automatic aircraft classification from radar measurements. The main advantage of this method is that radar measurements and subsequent classification can be achieved without the (active) cooperation of the aircraft (*Non-Cooperative Target Recognition*), which can potentially decrease the number of friendly fire incidents.

In this thesis aircraft are classified on the basis of *radar range profiles* which are measurements of an aircraft's radar reflectivity measured along the line-of-sight between the radar and the aircraft. Since the amount of reflected radar energy is different for different parts of the aircraft, a range is dependent on the geometry of the aircraft. It is this property of radar range profiles which make them suitable features for automatic aircraft recognition. However, the use of range profile for aircraft classification does have drawbacks, which are addressed in this thesis.

Radar range profiles are heavily influenced by the distance between the radar and the target at the time of measurement: changes in distance cause (cyclic) shifts of the range profile. The distance between the radar and the target can not be estimated with sufficient accuracy to correct the range profile for this effect. Therefore, a classifier for range profiles should be translation invariant, i.e., independent of any cyclic shift of the range profile.

Chapter 3 introduces a new method for obtaining translation invariant classification of range profiles. First it is shown that range profile alignment is essentially a problem of phase estimation. On the basis of this analysis, a translation invariant representation called the *Zero-Phase Representation* is constructed. The main advantage of this representation is that translation invariant classification can be achieved more rapidly than with existing methods, at the cost of a decrease in classification accuracy.

A second drawback of using range profiles for aircraft classification is the strong dependency of range profiles on the pose of the aircraft with respect to the radar (called *aspect angle*). Aircraft rotations cause a change in aspect angle, and even small rotations can have a drastic effect on the resulting range profile due to interference effects.

Although aspect angles can be estimated from *tracking* data (a sequence of three-dimensional aircraft positions), the accuracy of these estimates is rather low. Chapter 4 describes a new method for increasing these aspect angle estimates by analysing phase information extracted from the measured range profiles.

Perhaps the main drawback of using range profiles for aircraft classification is the large amount of training profiles required to construct a reliable classifier. Range profiles are usually represented as vectors containing hundreds of elements. It is well known that constructing sta-

tistical classifiers in high-dimensional vector spaces requires large quantities of training data. Furthermore, the dependency of range profiles on aspect angle causes large within-class variances, and so again many profiles are required to accurately reflect their statistical distribution. Unfortunately, measuring radar range profiles is both time-consuming and expensive, and so using measured data is not a feasible option for constructing a classifier.

Simulated range profiles provide an alternative source for obtaining training data. Using radar simulation software and CAD models of aircraft, large numbers of range profiles can be obtained cheaply and relatively quickly, and at any desired aspect angle.

Simulated range profiles are different from measured range profiles in a number of ways. Due to the limited accuracy of both the simulation software as well as the CAD models, simulated range profiles can only approximate measured range profiles. Furthermore, simulated range profiles are not corrupted by measurement noise. Therefore, when constructing a classifier for measured range profiles on the basis of a simulated training set, care must be taken to train the classifier solely on those features which are present in both simulated and measured range profiles.

In this thesis it is assumed that the common features present in both simulated and measured range profiles are the positions of local amplitude maxima of the profiles. These maxima are related to the position of string radar reflectors on the aircraft, and one can assume that at least some of the reflectors are present in both simulated and measured profiles.

Chapter 5 introduces a *generative model* of the statistical distribution of these common features, which is conditioned on aircraft type and aspect angle. The parameters of this model have a clear physical interpretation, and in chapter 5 it is shown how to reliably estimate the model parameters from a data set containing simulated range profiles.

Finally, chapter 6 shows how the estimated models can be used to classify (sequences of) *measured* range profiles. The experiments performed in this chapter show that the classification performance obtained is comparable to the performance of standard classifiers. The main advantage of the method described here is the clear physical interpretation of the model parameters, and the ability to estimate the statistical distribution of the common features at any desired aspect angle.

Samenvatting

De mogelijkheid om snel en betrouwbaar vliegtuigen te kunnen herkennen is van groot belang voor zowel civiele als militaire luchtvaart. Verkeersleiders dienen continu op de hoogte te zijn van de positie, snelheid en andere kenmerken van vliegtuigen die zich in het lokale luchtruim bevinden. In militaire scenario's is betrouwbare informatie omtrent vliegtuigtypes letterlijk van levensbelang.

De bestaande systemen voor het automatisch bepalen van het type vliegtuig (ofwel vliegtuig *classificatie*) zijn gebaseerd op een vraag-antwoord principe: een grondstation stuurt een zogenaamde *challenge* naar een vliegtuig, dat vervolgens geacht wordt een bijpassend antwoord terug te sturen. Een vliegtuig dat geen passend antwoord terugstuurd wordt dan geacht vijandig te zijn.

Helaas is dit laatste niet altijd het geval. In de recente geschiedenis zijn er meerder gevallen bekend waarin een vriendelijk vliegtuig niet in staat bleek een passend antwoord te versturen ten gevolge van een menselijke fout of door een gebrek aan de hardware, met alle gevolgen vandien.

Het onderwerp van dit proefschrift is het automatisch herkennen van het type vliegtuig (ofwel vliegtuig *classificatie*) op basis van radarmetingen. Het grote voordeel van deze methode is dat in dit geval de meting en de daaropvolgende classificatie plaatsvinden zonder de (actieve) medewerking van het betreffende vliegtuig (*non-cooperative target recognition*). Hierdoor kunnen hopelijk in de toekomst veel ongelukken worden voorkomen.

In dit proefschrift is gekozen voor het gebruik van *radarafstandsprofielen* als basis voor classificatie. Radarafstandsprofielen (of kortweg profielen) zijn metingen van de sterkte van het gereflecteerde radarsignaal, gemeten langs de kijkrichting van de radar (zie 1.1). Omdat de sterkte van de refelectie verschillend is voor verschillende onderdelen van het vliegtuig zijn profielen afhankelijk van de vorm van het vliegtuig. Het is deze afhankelijkheid die profielen geschikt maakt als basis voor vliegtuigclassificatie.

Het gebruik van radarafstandsprofielen voor vliegtuigclassificatie heeft echter ook nadelen, waarvan de belangrijkste in dit proefschrift besproken en opgelost worden.

Radarafstandsprofielen zijn sterk afhankelijk van de afstand tussen de radar en het vliegtuig tijdens de meting; een verandering in afstand veroorzaakt een (cyclische) verschuiving van het gemeten profiel. Omdat de afstand tussen radar en vliegtuig niet voldoende nauwkeurig kan worden bepaald om voor dit effect te corrigeren, is het noodzakelijk dat de gebruikte classificatie methode invariant is onder willekeurig cyclische verschuivingen van het gemeten profiel.

Hoofdstuk 3 beschrijft een nieuwe methode voor het translatie-invariant classificeren van radarafstandsprofielen. Eerst wordt aangetoond dat het uitlijnen van profielen in essentie een fase-schattingsprobleem is. Op basis van deze analyse wordt vervolgens een nieuwe translatie-invariante representatie van profielen gedefinieerd, de *Zero-Phase Represenatation*. Het grote voordeel van deze representatie is dat translatie-invariante classificatie aanmerkelijk sneller kan worden gerealiseerd dan met bestaande methodes, hoewel dit wel ten koste gaat van het uiteindelijke classificatieresultaat.

Een tweede nadeel van het gebruik van radarafstandsprofielen voor vliegtuigclassificatie is hun sterke afhankelijkheid van de *aangezichtshoek* waaronder de radar het vliegtuig waarneemt. Rotaties van het vliegtuig veroorzaken veranderingen in aangezihtshoek, en relatief kleine rotaties (in de orde van een tiende graad) kunnen al een grote invloed hebben op het

gemeten profiel als gevolg van interferenties.

Aangezichtshoeken kunnen worden geschat uit zogenaamde *tracking* data (een serie metingen van de positie van het vliegtuig). Echter, deze schatting is weinig nauwkeurig. Hoofdstuk 4 beschrijft een nieuwe methode om op basis van de oorspronkelijke schattingen van de aangezichtshoek en faseinformatie verkregen uit de geobserveerde radarafstandsprofielen een verbeterde schatting van de aangezichtshoek te construeren.

Het grootste nadeel van het gebruik van profielen voor vliegtuigclassificatie is de grote hoeveelheid profielen die nodig zijn om een betrouwbare classificator te construeren. Profielen worden gewoonlijk gerepresenteerd als vectoren in een hoog-dimensionale vectorruimte (meestal van een paar honderd tot enkele duizenden dimensies). Het is algemeen bekend dat het construeren van statistische classificatoren in zulke ruimtes bijzonder veel metingen vereist. Verder is, vanwege de grote afhankelijkheid van profielen voor de aangezichtshoek, de statistische variantie van een set profielen van één vliegtuig bijzonder groot, en zijn er dus zeer veel profielen vereist om deze variantie goed te kunnen beschrijven. Het meten van profielen is echter kostbaar en tijdrovend, waardoor het gebruik van gemeten profielen voor het construeren van een classificator geen reële optie is.

Een oplossing voor dit probleem is het gebruik van gesimuleerde profielen voor het verkrijgen van een dataset van voldoende grootte. Gesimuleerde profielen worden verkregen met behulp van radarsimulatiesoftware in combinate met drie-dimensionale CAD modellen van vliegtuigen. Op deze manier kunnen snel en goedkoop veel profielen worden berekend, bij elke gewenste aangezichtshoek, en voor elk gewenst vliegtuigtype.

Gesimuleerde radarafstandsprofielen zijn verschillen in een aantal opzichten van hun gemeten tegenhangers. Vanwege de beperkte nauwkeurigheid van zowel de simulatesoftware als de gebruikte CAD modellen zal een gesimuleerd profiel nooit exact gelijk zijn aan een gemeten profiel van hetzelfde vliegtuig bij dezelfde aangezichtshoek. Tevens zijn gesimuleerde profielen niet onderhevig aan ruis. Het gebruik van gesimuleerde profielen voor het classificeren van gemeten profielen zal dan ook gebaseerd moeten zijn enkel op die kenmerken die in beide types voorkomen.

Een belangrijke aanname in dit proefschrift is dat de 'gezamenlijke' kenmerken in gesimuleerde en gemeten profielen de positie van lokale maxima van de amplitudes in een profiel zijn. Deze maxima komen overeen met sterke reflectoren op een vliegtuig, en men mag verwachten dat op zijn minst een aantal van die reflectoren zowel in de simulatie als tijdens een echte meting in het profiel aanwezig zijn.

Hoofdstuk 5 introduceert een *generatief model* voor deze gezamenlijke kenmerken. Dit generatief model is in staat om voor een gegeven vliegtuig en aangezichtshoek de statistische distributie over de posities van lokale amplitude maxima te genereren. Bovendien bevat het model parameters die de effecten van rotaties van het vliegtuig op de kenmerken modelleren. Hoofdstuk 5 beschrijft hoe de modelparameters betrouwbaar geschat kunnen worden op basis van een dataset bestaande uit gesimuleerde profielen.

Hoofdstuk 6 laat vervolgens zien hoe de geschatte modellen gebruikt kunnen worden om *gemeten* profielen te classificeren. Uit de experimenten blijkt dat het uiteindelijke classificatieresultaat vergelijkbaar is met dat van standaardmethodes. De voordelen van de in dit proefschrift beschreven methode zijn dat het statistisch model gebaseerd is op fysisch interpretabele parameters, en dat het niet meer nodig is om een grote hoeveelheid profielen te gebruiken tijdens classificatie: het model is in staat om bij elke willekeurige aangezichtshoek voor elk gewenst vliegtuigtype de distributie van de gezamenlijke kenmerken te voorspellen .

Dankwoord

Velen hebben, direct dan wel indirect, bijgedragen aan de totstandkoming van dit proefschrift.

Frans Groen was in zijn rol als promotor nauw betrokken bij dit onderzoek. Zijn bijzonder grote kennis van zaken, in combinatie met zijn plezier in het spelen van advocaat van de duivel, maakten elke bespreking weer spannend. Tegelijkertijd zorgde zijn enthousiasme en optimisme ervoor dat ik altijd vrolijker uit een bespreking kwam dan dat ik er aan begon.

Ben Kröse heeft als geen ander invloed gehad op de loop van dit onderzoek en de inhoud van dit proefschrift. Zijn oog voor detail en talent voor het schrijven van wetenschappelijke teksten hebben in grote mate bijgedragen aan de kwaliteit van dit proefschrift.

Sjoerd Gelsema, verbonden aan TNO Fysisch en Elektronisch Laboratorium, was altijd bereid om zijn kennis van radarsystemen en signaalverwerking met mij te delen. Ook zijn adviezen met betrekking tot schrijfstijl hebben de kwaliteit van dit proefschrift positief beinvloed.

Het proefschrift van René van der Heiden, nu werkzaam bij NATO Consultation, Command and Control Agency, is in heel concrete zin het startpunt van mijn onderzoek geweest. Zijn enthousiasme voor dit onderzoek was zeer aanstekelijk, en als onuitputtelijke bron van interessante literatuur heeft hij een hoop overbodig werk weten te voorkomen.

De overige commissieleden, prof.dr. Adriaans, prof.dr. Hamran, prof.dr. Sloot, dr. Webb en prof.dr. Zwamborn, wil ik bedanken voor het beoordelen van dit proefschrift en het zitting nemen in mijn promotiecommissie.

De collega's van de Intelligente Autonome Systemen groep van de Universiteit van Amsterdam zorgden voor een plezierige werkomgeving. Met name de uitermate tolerante houding van mijn kamergenoten Stephan ten Hagen, Nikos Massios en Josep Porta met betrekking tot mijn soms wat eigenaardige muzikale voorkeuren heb ik bijzonder gewaardeerd.

Een deel van mijn onderzoek heb ik uitgevoerd in de groep Radarsystemen en Informatie-extractie van TNO-FEL. Ook daar vond ik een gezellige groep collega's, die nooit te beroerd waren mijn vragen te beantwoorden. De hulp van Lucas van Ewijk was onontbeerlijk bij het doorgronden van RAPPORT.

Zoals menig AIO kwam ook ik aan het eind van de rit in tijdnood. Ik ben dan ook zeer dankbaar voor de flexibiliteit van mijn nieuwe werkgever Chess iT.

Frank Sterk is verantwoordelijk voor de prachtige omslag, en heeft hiermee een wonderschone hattrick voltooid.

Ook buiten werktijd kon ik rekenen op veel morele steun, alsook de broodnodige afleiding. Met Camiel, Fergal, Kees en Tim is het al zo'n 15 jaar gezellig, en ongetwijfeld over 15 jaar nog steeds. Met Kim, Nanette, Tjals en Heinz werd het vaak veel te laat. Het volgende een-na-laatste rondje is van mij! Esther, Freya, Gerbrand, Jochem, Joost, Kirsten, Marcia en Mats hielpen mij door middel van voedsel, drank en heel veel spelletjes mij regelmatig de promotieperikelen even vergeten.

Ilja en Laura, als ex-AIO's wisten jullie met woord, daad en drassi altijd de dingen weer in het juiste perspectief te plaatsen. Binnenkort kunnen jullie hopelijk met mij mee rijden!

Papa en mama, jullie onvoorwaardelijke steun was onmisbaar, en de basis van alles. **Thuis** is waar jullie zijn.

Amsterdam, 22 oktober 2003