Iris Pattern Matching using Score Normalisation Techniques

Nigel Sedgwick

Cambridge Algorithmica Limited
9 Oakdene
Beaconsfield
Buckinghamshire
United Kingdom HP9 2BZ

Tel: +44 (0)1494 678989
URL: http://www.camalg.co.uk
Fax: +44 (0)1494 678990
Email: ncs@camalg.co.uk

Overview of Presentation

1. Score Normalisation and Multi-Biometric Fusion
2. Biometric Gain against Impostors (BGI) [1,2]
3. Likelihood Ratios and Simple Bayesian Fusion
4. HD Normalisation for IrisCode Matching
5. Result A: Can HD Normalisation be Improved?
6. Fusion for Multi-Algorithmic IrisCodes
7. Result B: Does BGI Normalisation/Fusion do Better?
8. Iris Recognition and ICE Looking Forward …

References

Summary

Multi-Biometric Fusion

Multi-biometric is a term covering:

- multi-modal: using 2 or more different biometric modalities; eg: iris combined with fingerprint; face combined with hand geometry
- multi-instance: combining more than one separate instance of the same biometric modality; eg fingerprints from 2 or more different fingers; irises of both eyes
- multi-algorithmic: processing the same biometric sample with 2 or more feature analysis and/or pattern-matching algorithms, and combining the results
- multi-presentational: (somewhat different in nature) capturing the same biometric instance (eg a single fingerprint) more than once, to reduce image capture errors, and then combining, or selecting the best result

Score Normalisation

- Raw scores can be on arbitrary, device-dependent or algorithm-dependent scales.
- It is meaningless to combine scores from different arbitrary scales.
- Score normalisation applies an appropriate transformation to scores from each modality, instance or algorithm, so that all normalised scores are on the same scale.

- Probability ordered scales: high scores match better.
- Distance ordered scales: low scores match better.

- Scores closely related to linear likelihood ratios are usually best combined by multiplication.
- Scores closely related to log likelihood ratios are usually best combined by addition.

Score Normalisation and Multi-Biometric Fusion

- This talk is principally about multi-algorithmic fusion, ie improved matching of a single biometric sample.
- Aspects apply to multi-instance fusion, eg fusing matches of left and right irises.
- Improved technical performance arises from both, in terms of better trade-off between False Match Rate (FMR) and False Non-Match Rate (FNMR).
- Multi-modal, multi-instance, and perhaps multi-algorithmic, fusion provide greater resistance to biometric avoidance techniques (eg gummy fingerprint overlays).
- Multi-modal fusion, eg iris and fingerprint, also provides greater universality (works for more people), as does multi-instance fusion to some degree.
- Score normalisation is essential to multi-algorithmic fusion and multi-modal fusion; often it benefits multi-instance fusion.

BGI: Biometric Gain against Impostors

BGI is rather like hi-fi amplifier gain: just consider the ratio of the output to the input, for each biometric modality, instance or algorithm.

$$\text{BGI} = \frac{\text{Probability of being an impostor, given the biometric evidence too}}{\text{Probability of being an impostor, given only prior knowledge}}$$

Most of the time, a very good approximation to the BGI is the reciprocal of the Likelihood Ratio Genuine to Impostor (LRGI). This is used in many good pattern-matching algorithms in existing biometric devices.

$$\text{BGI} \sim \frac{1.0}{\text{LRGI}} = \frac{\text{Probability of seeing the evidence from an impostor}}{\text{Probability of seeing it from the expected genuine subject}}$$

Relationship between BGI and LRGI

Bayes Rule: `P(I|E).P(E) = P(E|I).P(I)`
Probs sum to 1.0: `P(I) + P(G) = 1.0`
`P(I|E) + P(G|E) = 1.0`
Weighted sum: `P(E) = P(E|G).P(G) + P(E|I).P(I)`

`I` = Subject is Impostor
`G` = Subject is Genuine
`E` = Evidence (biometric score)

We desire the *a posteriori* probability: `P(I|E)`

$$P(I|E) = \frac{P(E|I).P(I)}{P(E|G).P(G) + P(E|I).P(I)}$$

$$BGI(E) = \frac{P(I|E)}{P(I)} = \frac{1.0}{P(I) + P(G).[P(E|G)/P(E|I)]} = \frac{1.0}{P(I) + P(G).LRGI(E)}$$

But often, we do not know the *a priori* probabilities: `P(I), P(G)`

So, assuming *a priori* probability of Impostor is very small: $BGI(E) \sim \frac{1.0}{LRGI(E)}$

Bayesian Fusion is Exact using Full LRGI

The definition of the Likelihood Ratio Genuine to Impostor is:

$$\mathrm{LRGI}(E) = \frac{P(E|G)}{P(E|I)}$$

If the total biometric evidence comes from a multi-biometric fusion (of individual scores e1, e2, e3, etc), this can be represented (exactly) as:

$$\mathrm{LRGI}(E) = \frac{P(E|G)}{P(E|I)} = \frac{P(e1,e2,e3,...|G)}{P(e1,e2,e3,...|I)}$$

Note that, very importantly, the LRGI is independent of the *a priori* probabilities: `P(I)` and `P(G)`.

Thus multi-biometric fusion can be done, using LRGIs, without any need for knowledge of the *a priori* probabilities.

In so far as *a priori* probabilities are important (eg if `P(I)` is not very small), these can be applied (as on the previous slide) after multi-biometric fusion.

Simple Bayesian Fusion is usually only an Approximation (often a Very Good One)

- **The basic assumption of simple Bayesian fusion is that the biometric measurements (e1, e2, e3, …) are statistically independent.**

$$\mathrm{LRGI}(E) = \frac{P(e1,e2,e3,...|G)}{P(e1,e2,e3,...|I)} \sim \frac{P(e1|G)}{P(e1|I)} * \frac{P(e2|G)}{P(e2|I)} * \frac{P(e3|G)}{P(e3|I)} ...$$

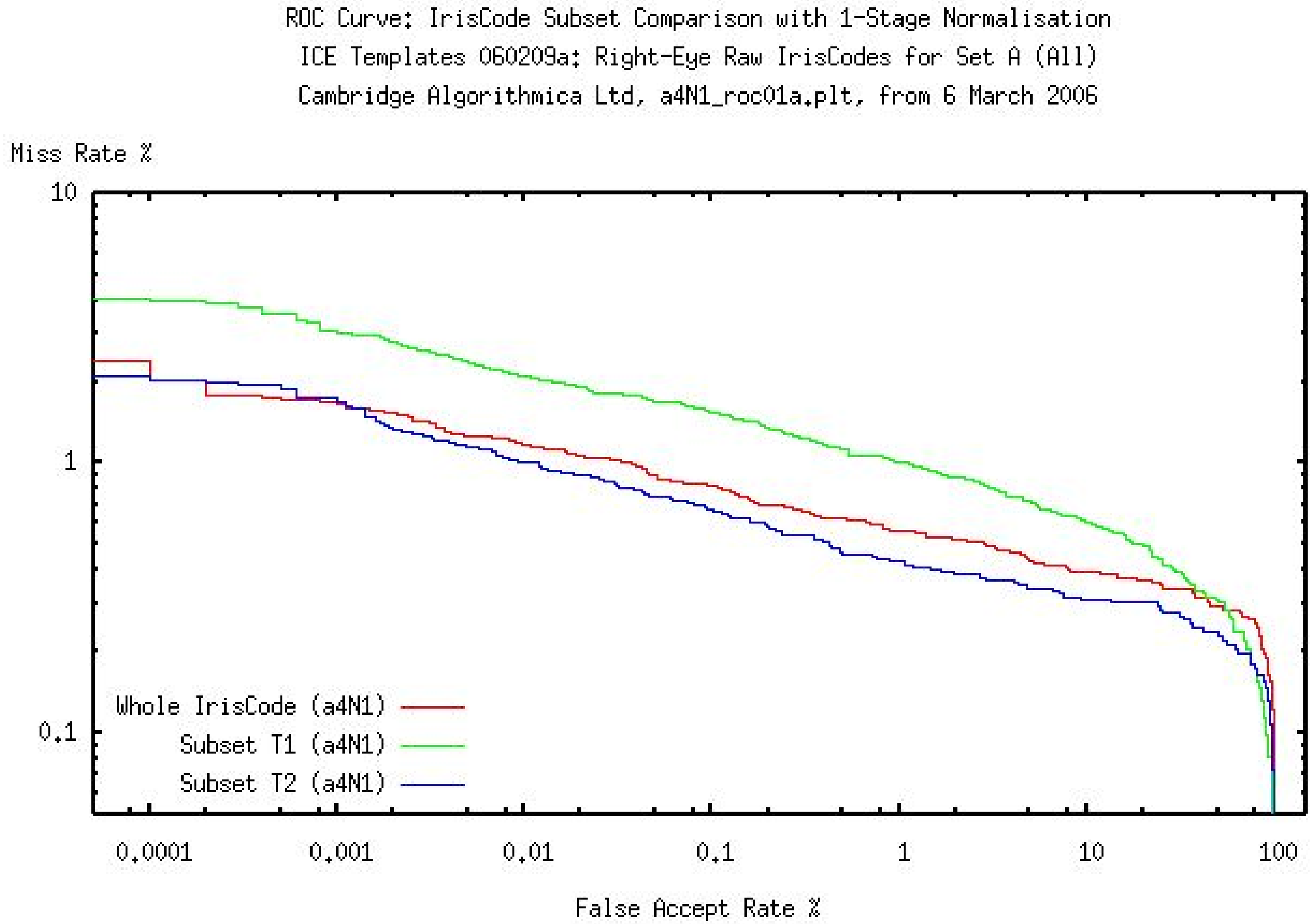
- **This is often not the case; it most certainly is not the case for multi-algorithmic fusion.**
- **However, experiments show [3] that pattern recognition algorithms based on simple Bayesian fusion are very often highly competitive with more complicated and sophisticated approaches.**
- **Furthermore, theory shows [4] that simple Bayesian fusion is optimal over wide ranges of the LRGI values to be combined, though not all.**
- **So, for these reasons, we view simple Bayesian fusion as an approach that should be evaluated for every application that calls for multi-biometric fusion.**

Notes on Modified IrisCodes and Multi-Algorithmic Definitions

- **Work reported here is only concerned with pattern matching of IrisCodes. No work is done on image processing.**
- **Here, IrisCodes differ from Daugman's original 2-bit quantised phasors [5].**
- **Each IrisCode is divided into 2 subsets: T1 and T2.**
- **Each subset is a different algorithmic analysis of the whole iris image.**
- **Initial pattern matching is done separately on each of the T1, T2 subsets, using the usual Hamming Distance approach.**
- **Initial pattern matching is also done treating the two subsets as a single entity. This is usually referred to as "whole IrisCode" or "all".**
- **The underlying direction of this work has been to investigate improved ways (multi-algorithmic fusion) of combining the analysis of the T1 and T2 subsets.**

Initial ROC Curve Comparison of T1 and T2 Subsets, and Whole IrisCode

Note unexpected ranking of T2 Subset, as better than whole IrisCode.

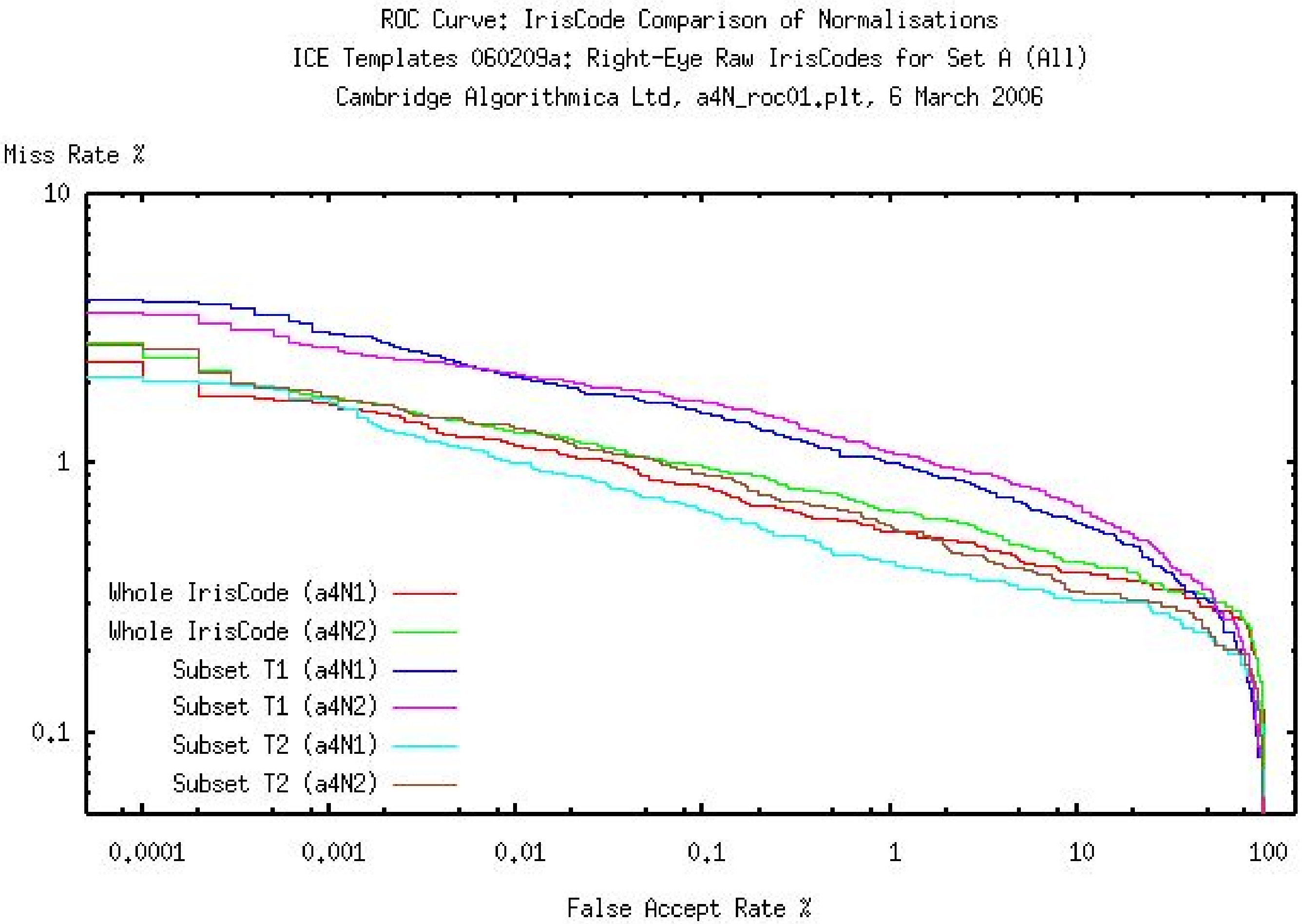


Hamming Distance Normalisation for IrisCode Matching

- Daugman's normalisation [6] is well-established and has two stages.
- This is to deal with variation, between iris image pairs, of the number of bits actually compared to form the Hamming Distance.
- **Stage 1:** the raw Hamming Distance (HD_{raw}) is given by the number of bits differing between the 2 IrisCodes divided by the number of bits compared (n, as determined from the probe and gallery mask bits).
- **Stage 2:** this modifies HD_{raw} non-linearly, leading to HD_{norm}.
- By mistake, but fortuitously, **our initial implementation of IrisCode pattern matching used only the Stage 1 Normalisation.**

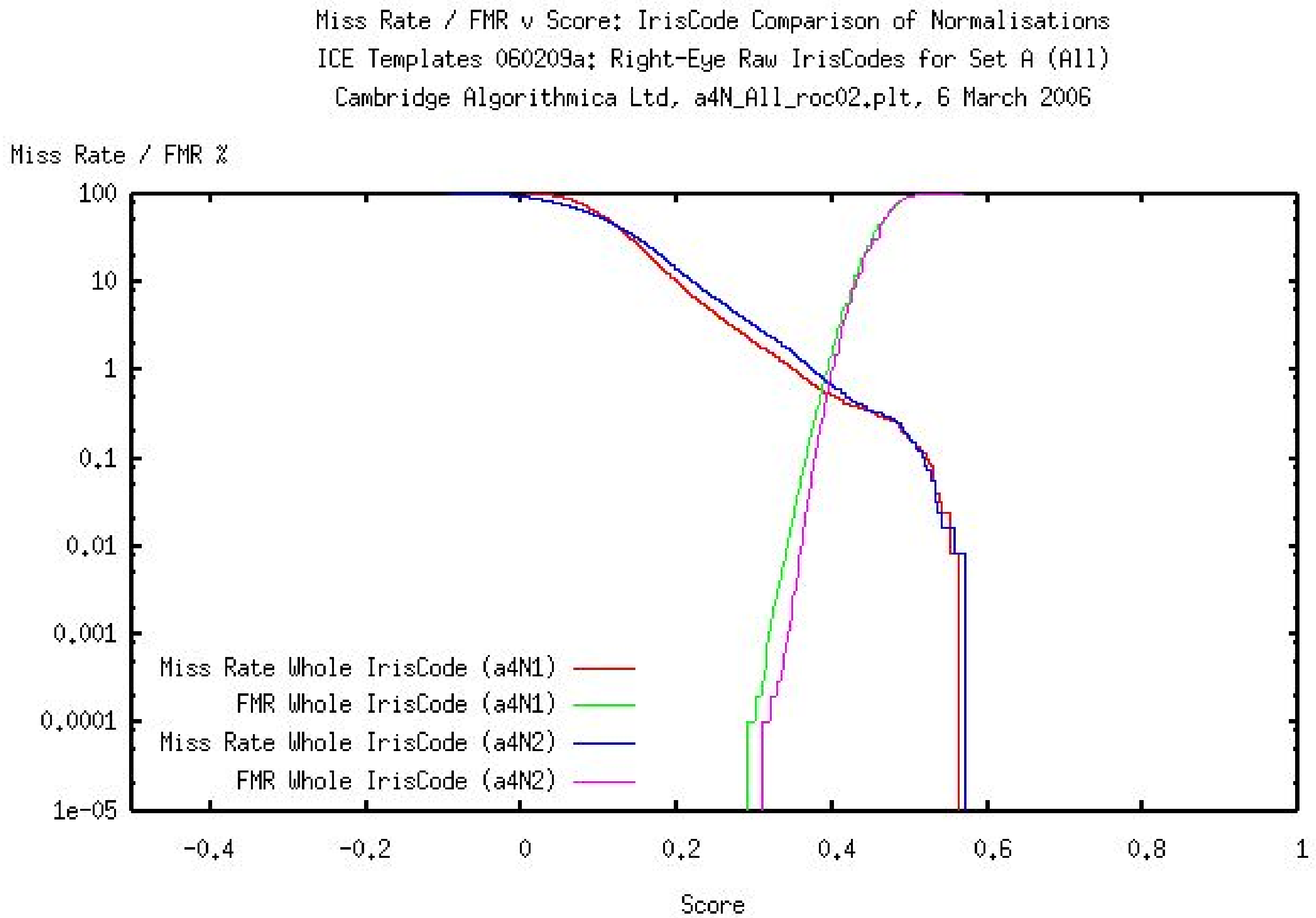
ROC Curve Comparison for 1-Stage (N1) and 2-Stage (N2) HD Normalisation

Result A: Unexpectedly, over most of each pair of curves (same T1, T2 or All), 1-stage normalisation is better than 2-stage normalisation.



Investigation into Why 1-Stage Normalisation is Better

Note: 2-stage normalisation is better for FMR; 1-stage normalisation is better for Miss Rate. Effect on Miss Rate mostly dominates, so 1-stage normalisation is mostly better.



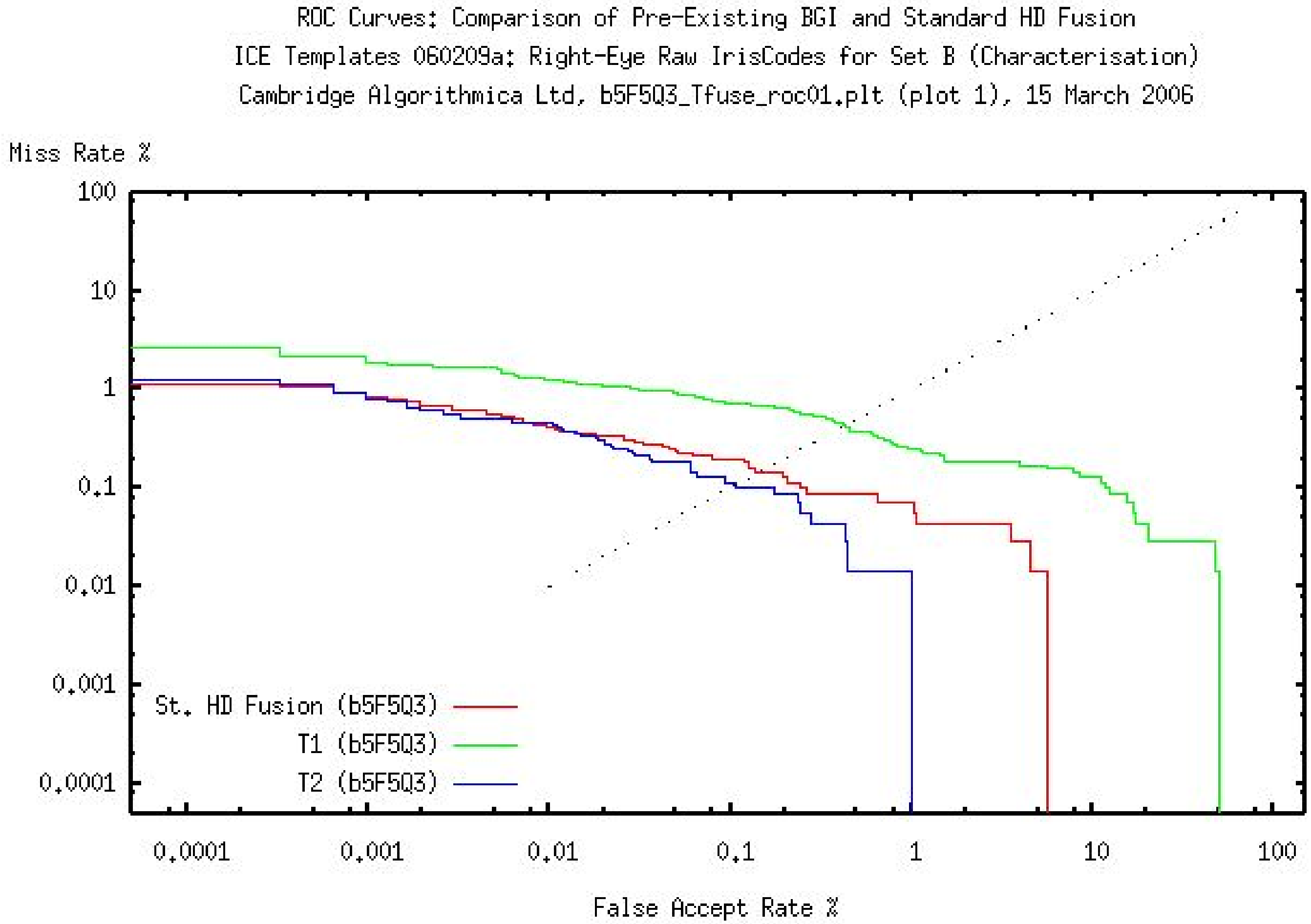
Note on Datasets Used

- **In this work, the BGI multi-algorithmic fusion uses some pattern matching results to characterise the PDFs of raw scores (HD_{raw}) for genuine matches and for impostor matches.**
- **It is appropriate to separate the evaluation dataset from the characterisation dataset.**
- **Accordingly, the IrisCodes from all right-iris images have been divided into 2 Sets.**

	Subjects	IrisCodes	Genuine Matches	Impostor Matches
Set A (B, C combined)	132	1,426	12,221	1,003,804
Set B (characterisation)	66	796	7,024	309,386
Set C (evaluation)	66	630	5,197	192,938
Sum: Set B + Set C	132	1,426	12,221	[502,324]

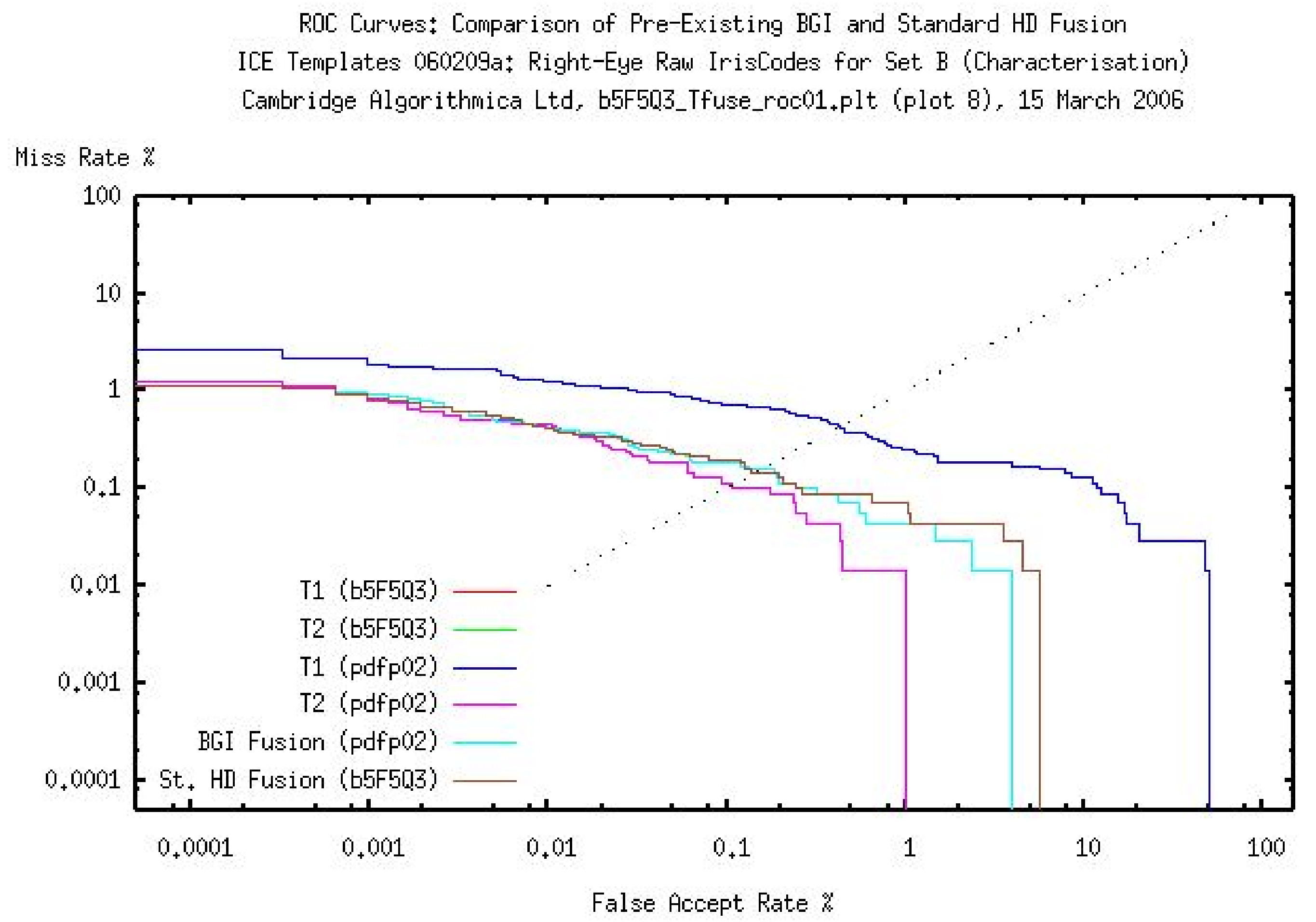
BGI Fusion: The Starting Point

ROC curves for T1 and T2 Subsets, and for All, each with 1-stage normalisation (and correction of the number of micro-rotations). Note re-labelling the “all” case as “Standard HD Fusion”.



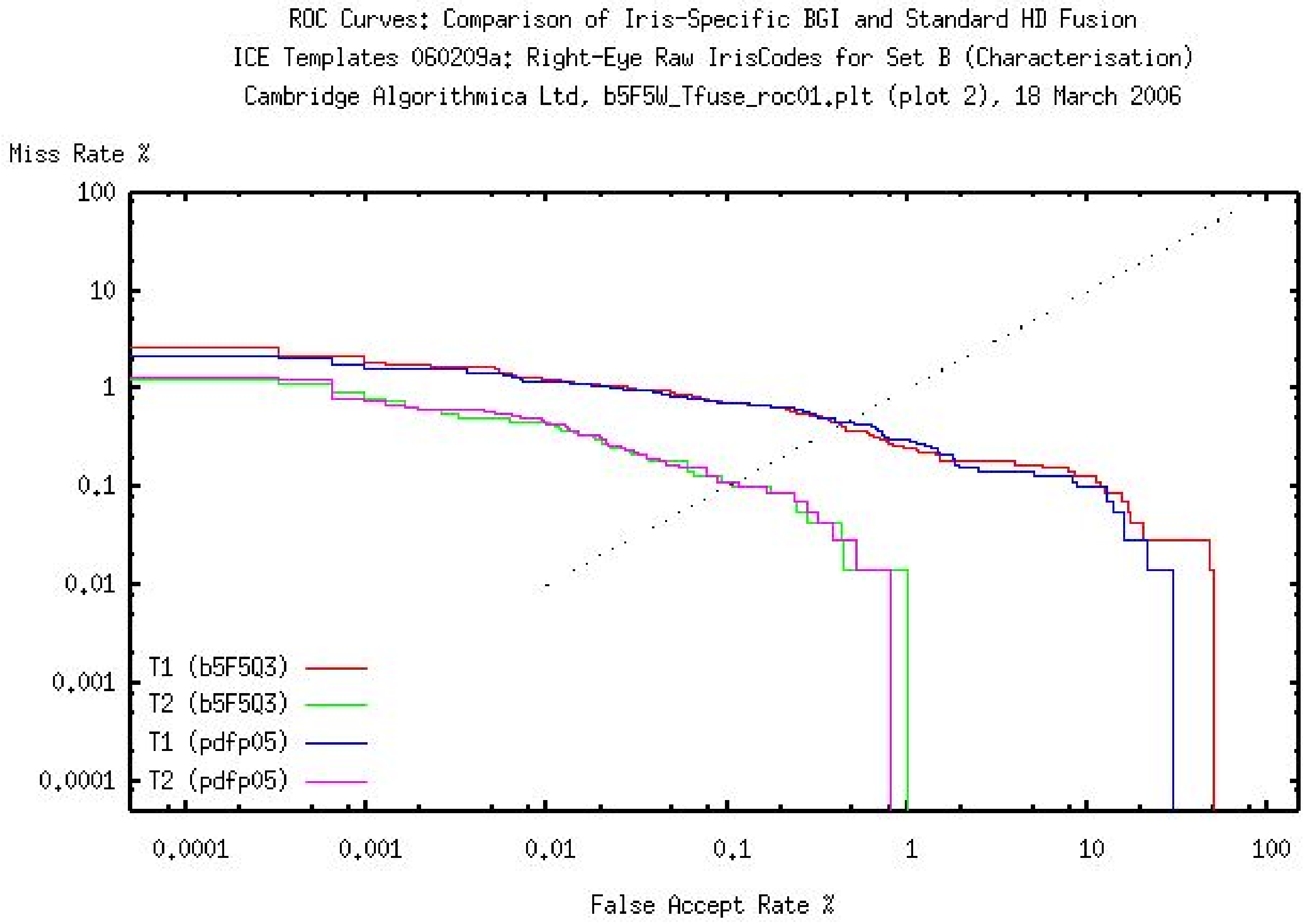
BGI Fusion: PDF Modelling using Best Pre-Existing Approach (from BSSR1 work)

Selected pre-existing PDF models (pdfp02, [2]) fail to do better than T2 Subset only. Performance very close to Standard HD fusion.



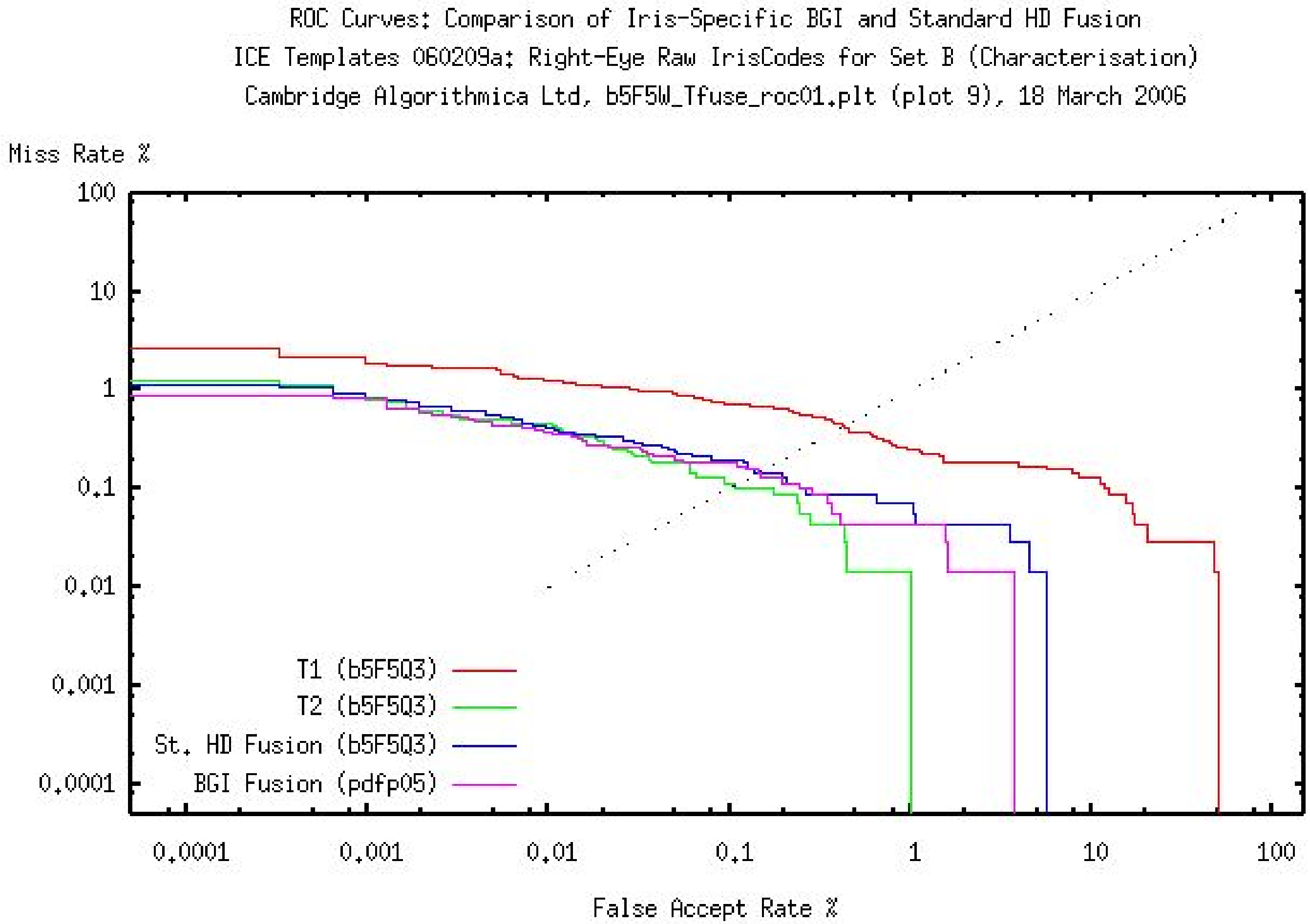
BGI Fusion: New IrisCode-Specific PDF Modelling

These PDF models (pdfp05) are shown here normalising T1 and T2 Subsets. These normalisations are not monotonic. Therefore the ROC curves before/after normalisation do not exactly overlay each other.



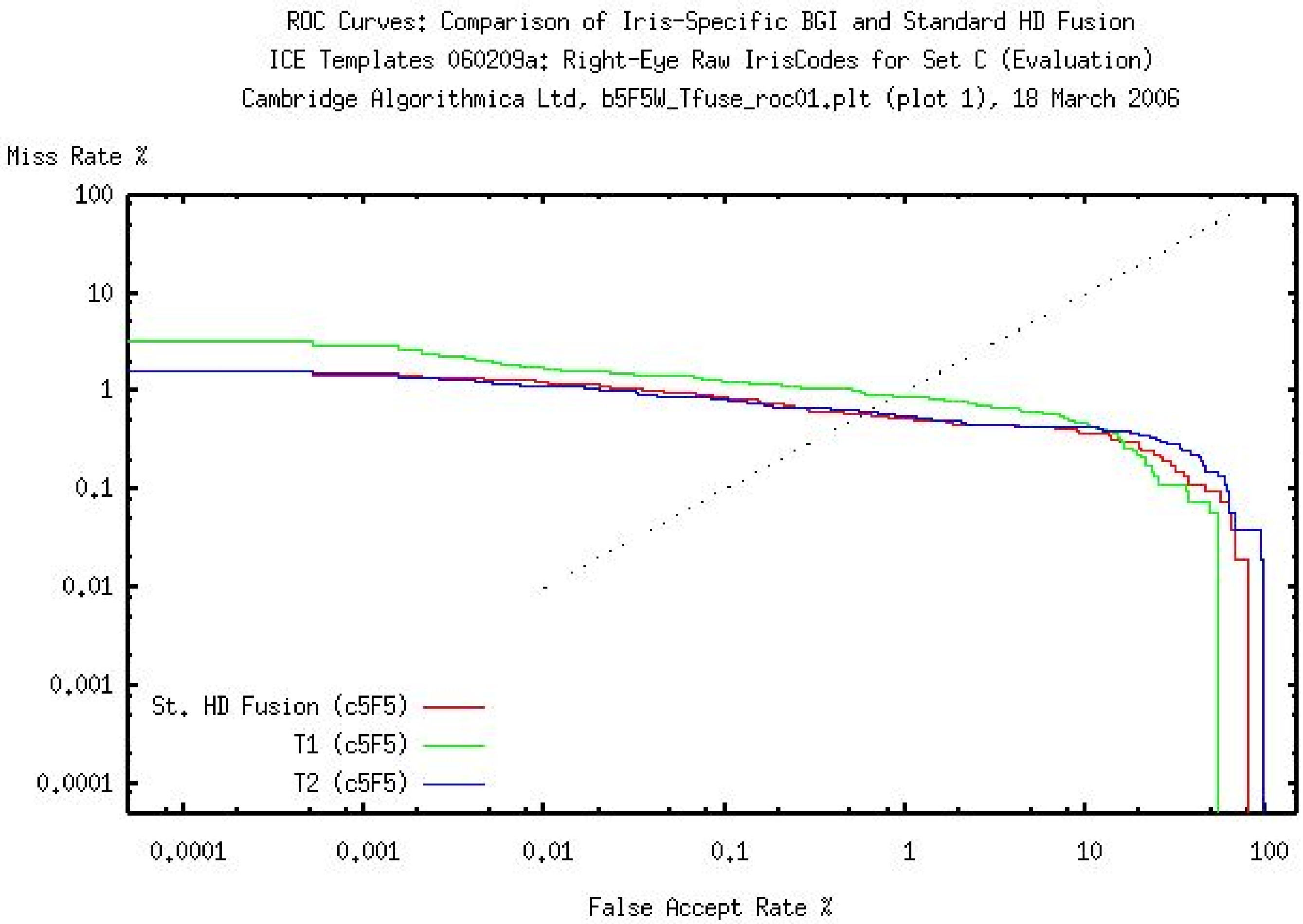
BGI Fusion: ROC Curve for Fusion after IrisCode-Specific Normalisation

These PDF models (pdfp05) also do not give better performance than the T2 Subset, just as with pdfp02 models and with Standard HD Fusion. [Note: not shown, pdfp05 models may do slightly better than pdfp02.]



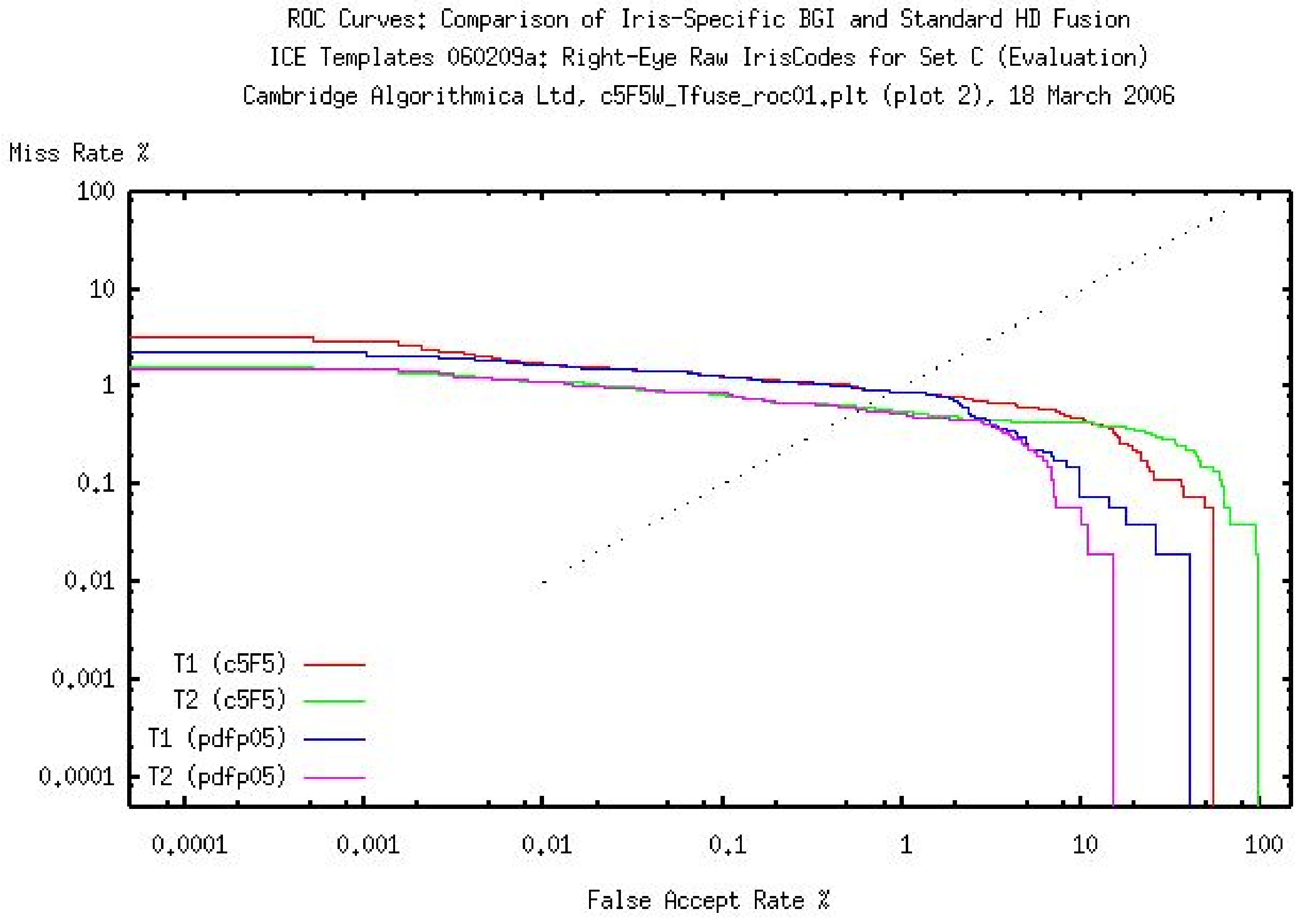
BGI Fusion: Starting Point ROC Curves for Evaluation Dataset (Set C)

These curves show different performance from the Characterisation Dataset (Set B). Overall performance is worse. The T2 Subset is no longer better than Standard HD Fusion. The T1 subset is better, over part of the range.



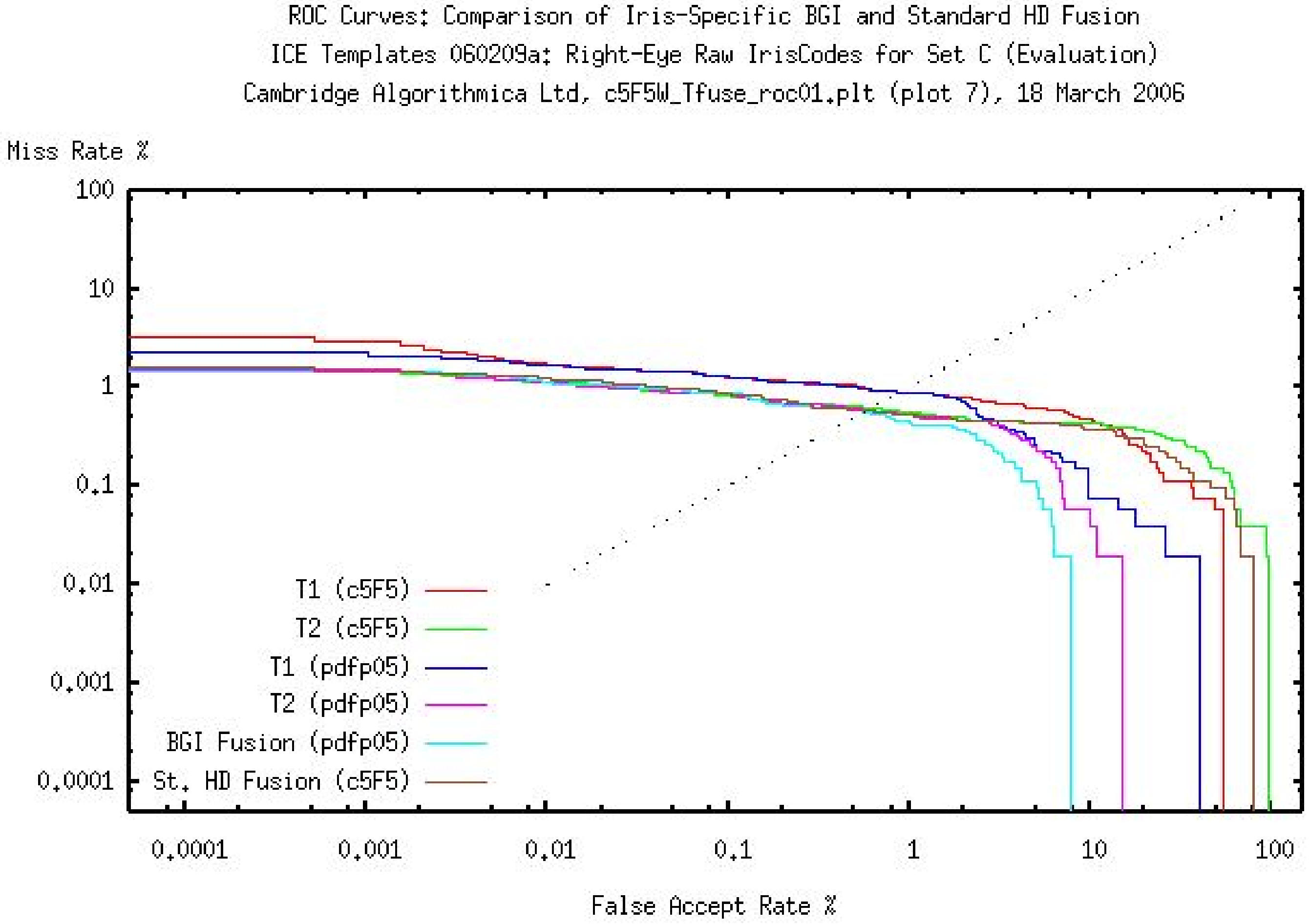
BGI Fusion: T1 and T2 Subset ROC Curves on Evaluation Dataset (Set C)

On the Evaluation Dataset, unlike for Set B, the pdfp05 normalised T1 and T2 Subsets show consistently better performance over some of the range.



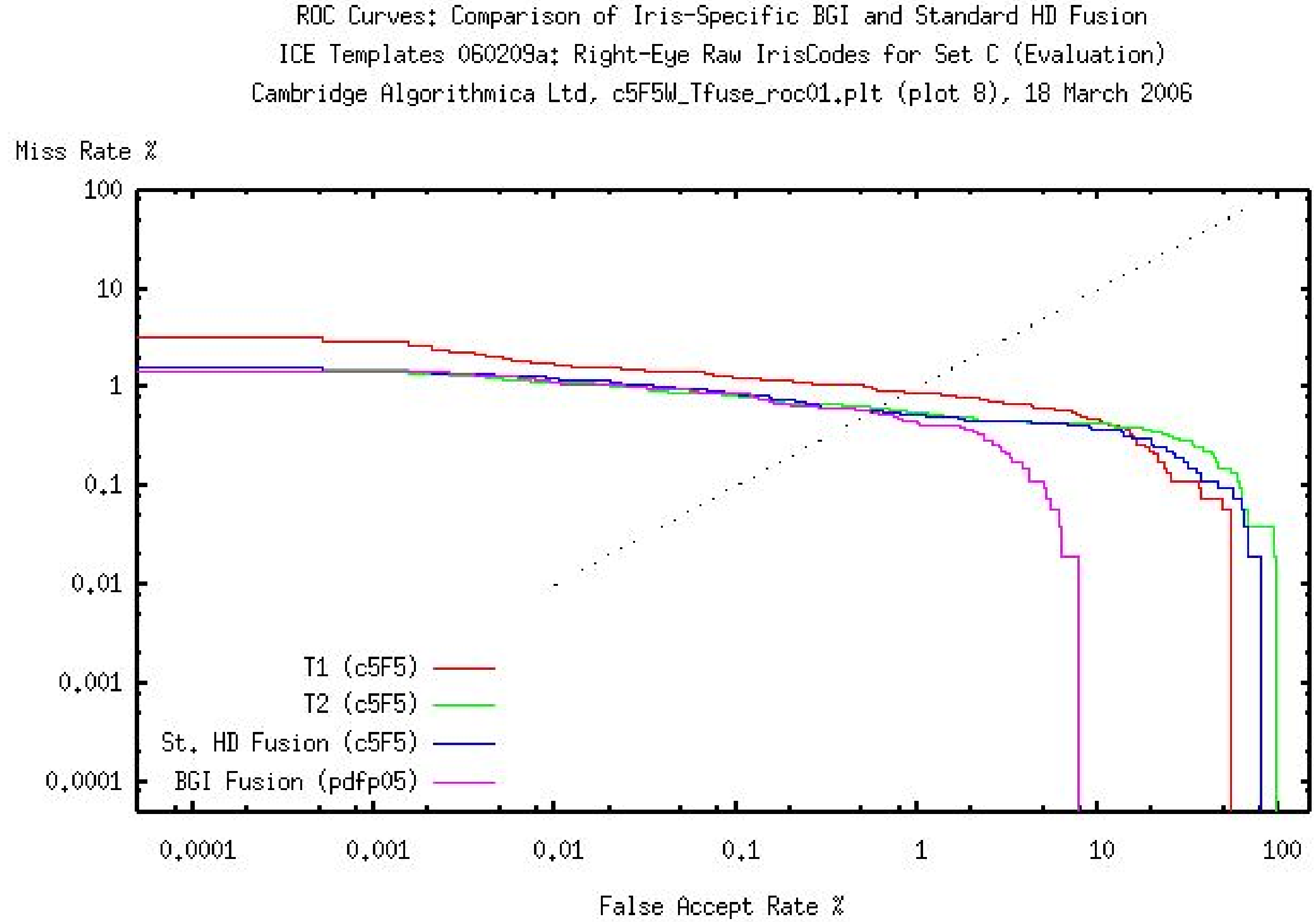
BGI Fusion: ROC Curves on Evaluation Dataset (Set C)

The Iris-Specific normalisation (pdfp05) clearly shows better performance over part of the range. Performance equals or exceeds, over the whole range, the better of the contributing multi-algorithmic features.



BGI Fusion: Final ROC Curves

Result B: BGI Iris-Specific Normalisation clearly shows better or equal performance over the whole ROC curve, on the Evaluation Dataset (Set C). Caveats: the datasets are not large; Set B gives conflicting results. The region of improvement is of most interest for multi-instance and multi-modal fusion.



References

[1] Nigel Sedgwick, ***The Need for Standardisation of Multi-Model Biometric Combination***, Cambridge Algorithmica Ltd, 6 November 2003, http://www.camalg.co.uk/s03017_pr0/s03017_pr0.pdf

[2] Nigel Sedgwick, ***Preliminary Report on Development and Evaluation of Multi-Biometric Fusion using the NIST BSSR1 517-Subject Dataset***, Cambridge Algorithmica Ltd, 28 May 2005, http://www.camalg.co.uk/s05011_tr0/s05011_tr0.pdf

[3] Pedro Domingos and Michael Pazzini, ***Beyond Independence: Conditions for the Optimality of the Simple Bayesian Classifier***, Proc 13th Int Conf on Machine Learning (ICML), 1996.

[4] Pedro Domingos and Michael Pazzini, ***On the Optimality of the Simple Bayesian Classifier under Zero-One Loss***, Machine Learning 1-30, 1997

[5] John Daugman, ***How Iris Recognition Works***, IEEE Trans on Circuits and Systems for Video Technology, CVST 14(1), January 2004, http://www.cl.cam.ac.uk/users/jgd1000/csvt.pdf

[6] John Daugman, ***Probing the Uniqueness and Randomness of IrisCodes: Results from 200 Billion Iris Pair Comparisons***, Cambridge University Computer Laboratory.

ICE 2006: Looking Forward

Thoughts on things that impact, potentially, on performance:

- **Left/Right Iris Multi-Instance Fusion:** provide iris pairs captured together or at the same time.
- **Possible Multi-Modal Use:** more emphasis on whole-ROC performance, rather than at a single operating point (eg 0.1% FAR).
- **Metadata on Sample Capture Equipment:** sample by sample details on manufacturer, model and version allows use of normalisations according to probe/gallery device pair.
- **Demography/Ethnicity of Gallery Subjects:** again, there are potential benefit for normalisation, using such knowledge as is available.

Summary

Theory of Normalisation and Fusion. **The BGI approach is based on normalisation of scores (from each contributing modality, instance or algorithm) to be likelihood ratios. Such fusion is independent of *a priori* probabilities. Simple Bayesian Fusion is a good approach, with theoretical and experimental support.**

Result A: **Iris recognition ROC performance with the modified IrisCodes is improved by using the 1-stage normalisation, rather than the 2-stage normalisation. [This should be investigated for the standard IrisCodes, as originally defined by Daugman.]**

Result B: **With the modified IrisCodes, using multi-algorithmic analysis of each whole iris image, an improvement in ROC performance was found using the BGI approach to normalisation and fusion. This was over the Standard Hamming Distance Fusion approach, of just treating all IrisCode bits together and deriving a single Hamming distance.**

Pursue BGI Approach. **It has definite potential to be useful. Caveats on work to date are: small datasets used for characterisation and evaluation; known poor images in characterisation dataset; conflicting results from the two datasets.**