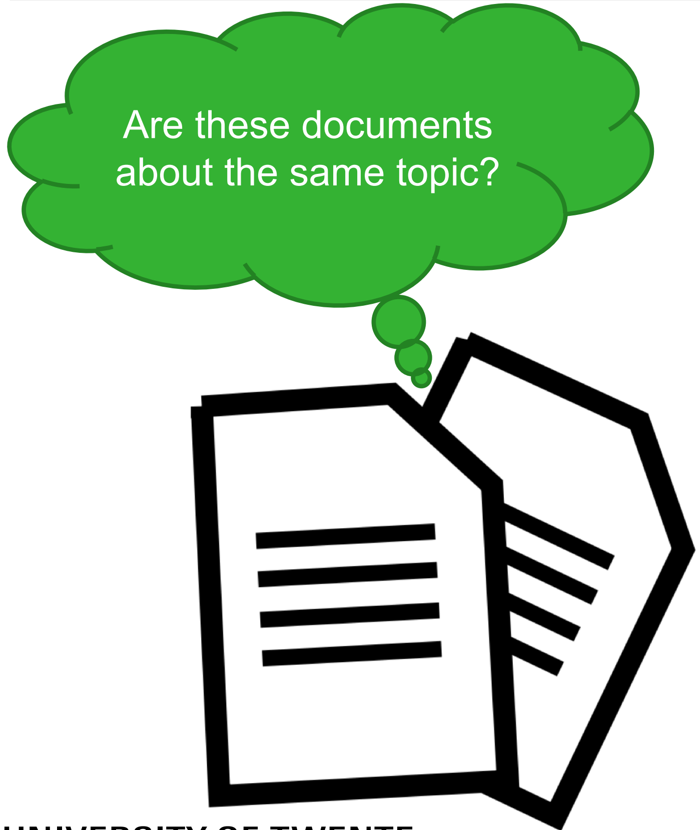
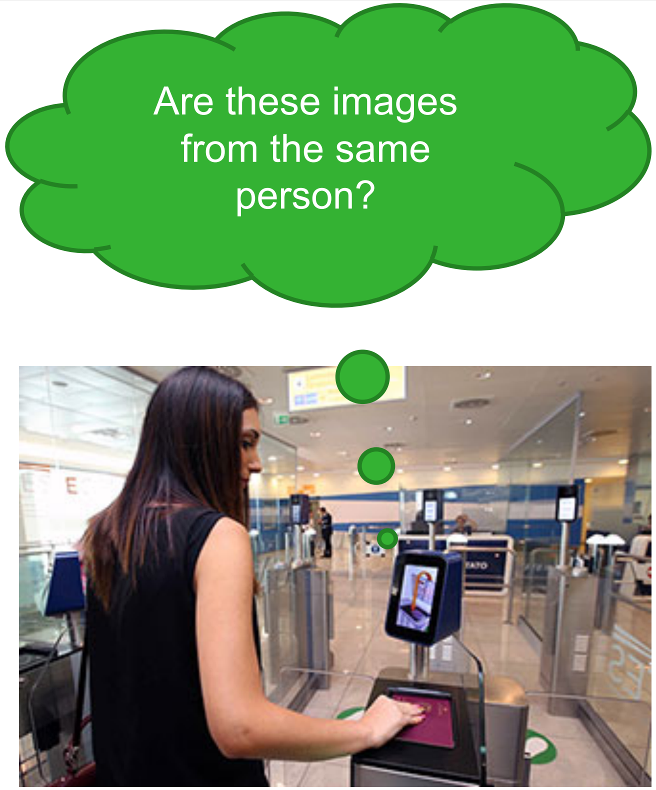
UNIVERSITY OF TWENTE.

A likelihood ratio classifier for histogram features

Raymond Veldhuis (UT), Kiran Raja (NTNU)

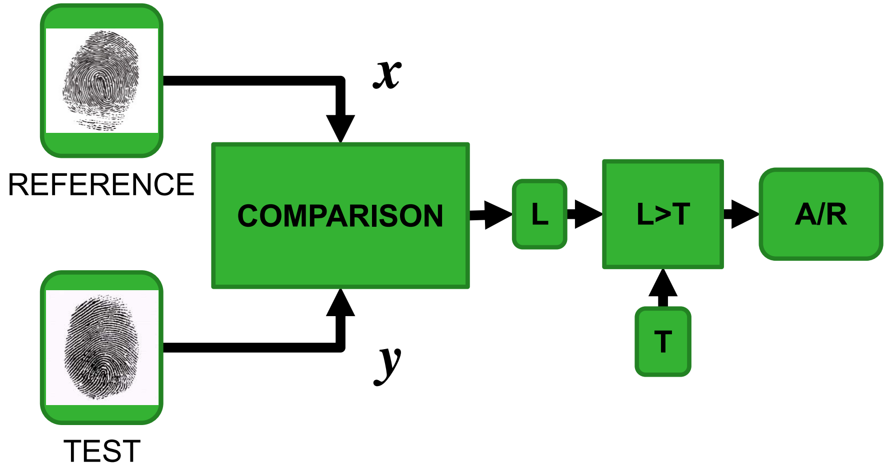
What is the question that the classifier must answer?



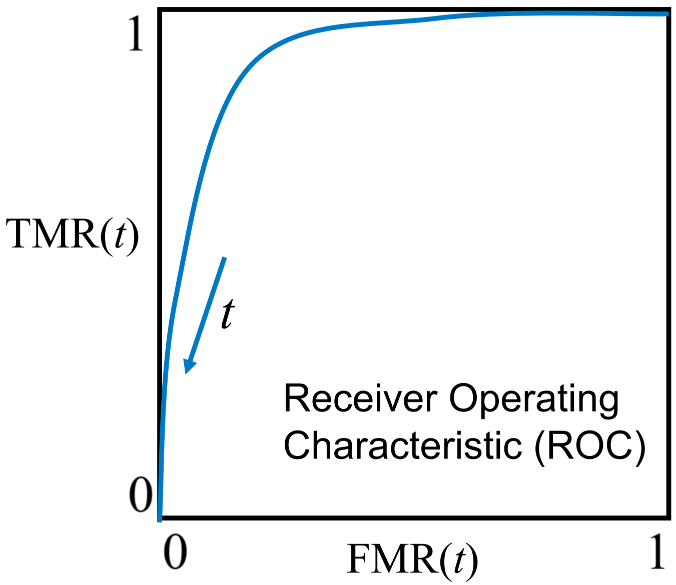


(BIOMETRIC) CLASSIFICATION THEORY

THE LR IS OPTIMAL IN NEWMAN-PEARSON SENSE



$$\mathrm{lr}(\boldsymbol{x},\boldsymbol{y}) \stackrel{\text{def}}{=} \frac{f_{\underline{\boldsymbol{x}},\underline{\boldsymbol{y}}}(\boldsymbol{x},\boldsymbol{y}|S)}{f_{\underline{\boldsymbol{x}},\underline{\boldsymbol{y}}}(\boldsymbol{x},\boldsymbol{y}|D)},$$



THE LIKELIHOOD RATIO

REWARDS RARENESS

$$\mathrm{lr}(\boldsymbol{x},\boldsymbol{y}) \stackrel{\mathrm{def}}{=} \frac{f_{\underline{\boldsymbol{x}},\underline{\boldsymbol{y}}}(\boldsymbol{x},\boldsymbol{y}|S)}{f_{\underline{\boldsymbol{x}},\underline{\boldsymbol{y}}}(\boldsymbol{x},\boldsymbol{y}|D)} = \frac{f_{\underline{\boldsymbol{x}},\underline{\boldsymbol{y}}}(\boldsymbol{x},\boldsymbol{y}|S)}{f_{\underline{\boldsymbol{x}}}(\boldsymbol{x})f_{\underline{\boldsymbol{y}}}(\boldsymbol{y})}$$

	Case 1		Case 2	
	trace	suspect	trace	suspect
Heights	175	176	200	201
LR				

THE LIKELIHOOD RATIO

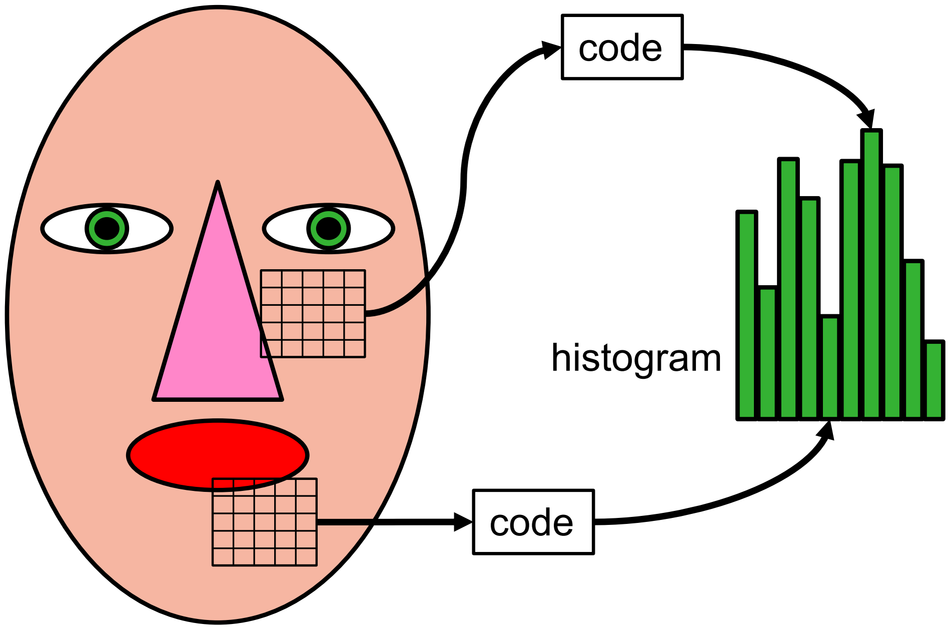
REWARDS RARENESS

$$\mathrm{lr}(\boldsymbol{x}, \boldsymbol{y}) \stackrel{\text{def}}{=} \frac{f_{\underline{\boldsymbol{x}},\underline{\boldsymbol{y}}}(\boldsymbol{x}, \boldsymbol{y}|S)}{f_{\underline{\boldsymbol{x}},\underline{\boldsymbol{y}}}(\boldsymbol{x}, \boldsymbol{y}|D)} = \frac{f_{\underline{\boldsymbol{x}},\underline{\boldsymbol{y}}}(\boldsymbol{x}, \boldsymbol{y}|S)}{f_{\underline{\boldsymbol{x}}}(\boldsymbol{x}) f_{\underline{\boldsymbol{y}}}(\boldsymbol{y})}$$

	Case 1		Case 2	
	trace	suspect	trace	suspect
Heights	175	176	200	201
LR	4		1170	

Histogram features

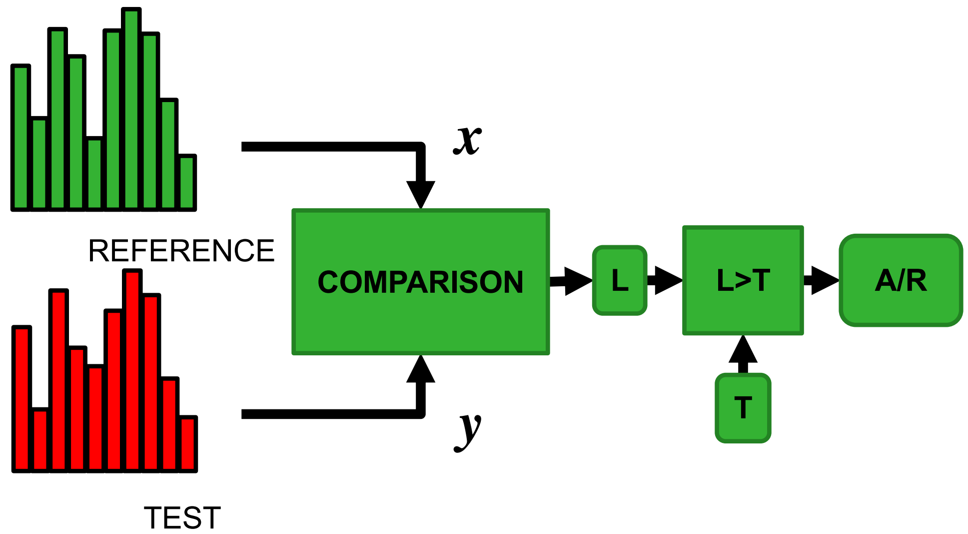
- Bag-of-words
  - Unigrams, n-grams
- Texture classifiers
  - LPB
  - BSIF
  - HOG
  - …



LR classifier for histogram comparison

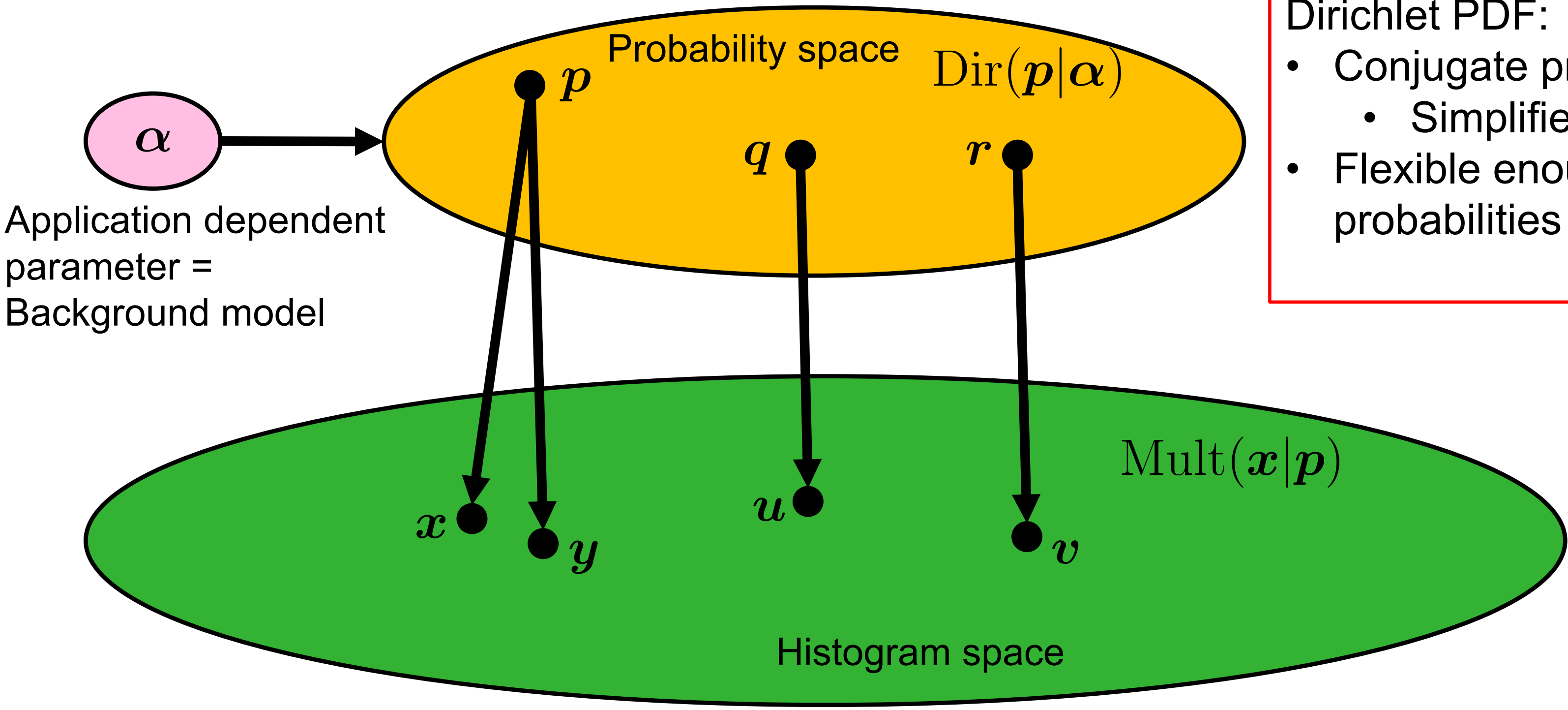
$$P\{\underline{\boldsymbol{x}} = \boldsymbol{x}|\boldsymbol{p}\} = \mathrm{Mult}(\boldsymbol{x}|\boldsymbol{p}) \stackrel{\mathrm{def}}{=} \frac{X!}{\prod_{i=1}^{n} x_i!}\prod_{i=1}^{n} p_i^{x_i}$$

$$P\{\underline{\boldsymbol{y}} = \boldsymbol{y}|\boldsymbol{q}\} = \mathrm{Mult}(\boldsymbol{y}|\boldsymbol{q})$$



Same: Histograms share the same bin probabilities
Different: Histograms have different bin probabilities

Histogram generation process



Dirichlet PDF:

- Conjugate prior of multinomial
  - Simplifies derivations
- Flexible enough to model bin probabilities

LR for histogram comparison

$$\mathrm{lr}(\boldsymbol{x}, \boldsymbol{y}|\boldsymbol{\alpha}) = \frac{\int_{\boldsymbol{p}} \mathrm{Mult}(\boldsymbol{x}|\boldsymbol{p})\mathrm{Mult}(\boldsymbol{y}|\boldsymbol{p})\mathrm{Dir}(\boldsymbol{p}|\boldsymbol{\alpha})\mathrm{d}\boldsymbol{p}}{\int_{\boldsymbol{p}} \mathrm{Mult}(\boldsymbol{x}|\boldsymbol{p})\mathrm{Dir}(\boldsymbol{p}|\boldsymbol{\alpha})\mathrm{d}\boldsymbol{p} \int_{\boldsymbol{q}} \mathrm{Mult}(\boldsymbol{y}|\boldsymbol{q})\mathrm{Dir}(\boldsymbol{q}|\boldsymbol{\alpha})\mathrm{d}\boldsymbol{q}}.$$

Dirichlet PDF with parameters **α** (background model)

Gamma function

$$\mathrm{Dir}(\boldsymbol{p}|\boldsymbol{\alpha}) \stackrel{\mathrm{def}}{=} \frac{\Gamma\left(\sum_{i=1}^{n} \alpha_i\right)}{\prod_{i=1}^{n} \Gamma(\alpha_i)} \prod_{i=1}^{n} p_i^{\alpha_i - 1}$$

$$= \frac{1}{\mathrm{B}(\boldsymbol{\alpha})} \prod_{i=1}^{n} p_i^{\alpha_i - 1}$$

Beta function

Dirichlet PDF:
- Conjugate prior of multinomial
  - Simplifies derivations
- Flexible enough to model bin probabilities

LR for histogram comparison - result

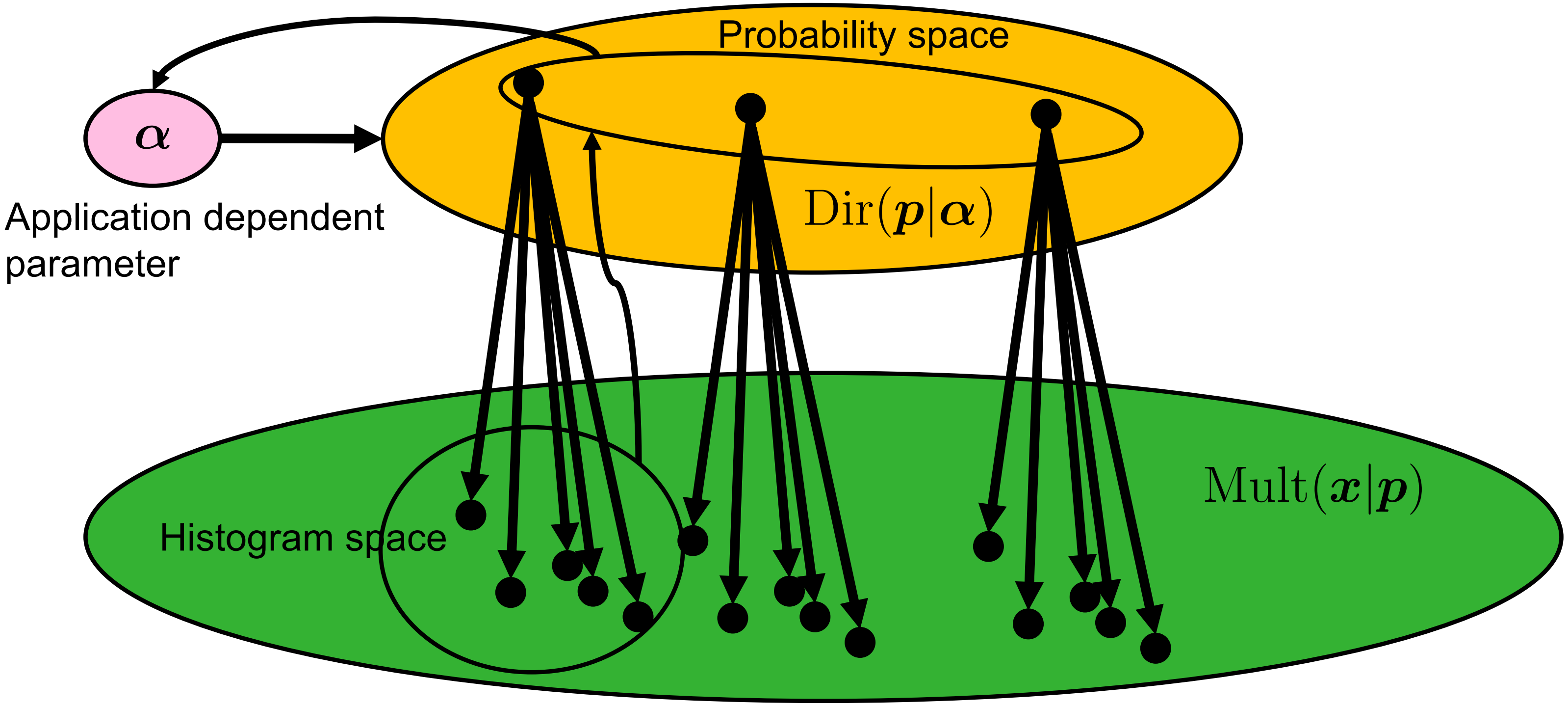
$$\mathrm{lr}(\boldsymbol{x}, \boldsymbol{y}|\boldsymbol{\alpha}) = \frac{\int_{\boldsymbol{p}} \mathrm{Mult}(\boldsymbol{x}|\boldsymbol{p})\mathrm{Mult}(\boldsymbol{y}|\boldsymbol{p})\mathrm{Dir}(\boldsymbol{p}|\boldsymbol{\alpha})\mathrm{d}\boldsymbol{p}}{\int_{\boldsymbol{p}} \mathrm{Mult}(\boldsymbol{x}|\boldsymbol{p})\mathrm{Dir}(\boldsymbol{p}|\boldsymbol{\alpha})\mathrm{d}\boldsymbol{p} \int_{\boldsymbol{q}} \mathrm{Mult}(\boldsymbol{y}|\boldsymbol{q})\mathrm{Dir}(\boldsymbol{q}|\boldsymbol{\alpha})\mathrm{d}\boldsymbol{q}}$$

$$= \mathrm{B}(\boldsymbol{\alpha})\frac{\mathrm{B}(\boldsymbol{x}+\boldsymbol{y}+\boldsymbol{\alpha})}{\mathrm{B}(\boldsymbol{x}+\boldsymbol{\alpha})\mathrm{B}(\boldsymbol{y}+\boldsymbol{\alpha})}.$$

Efficient computation in log-domain

Remaining problem: estimation of $\boldsymbol{\alpha}$ from training data

Histogram generation process



ML estimation of Dirichlet parameters

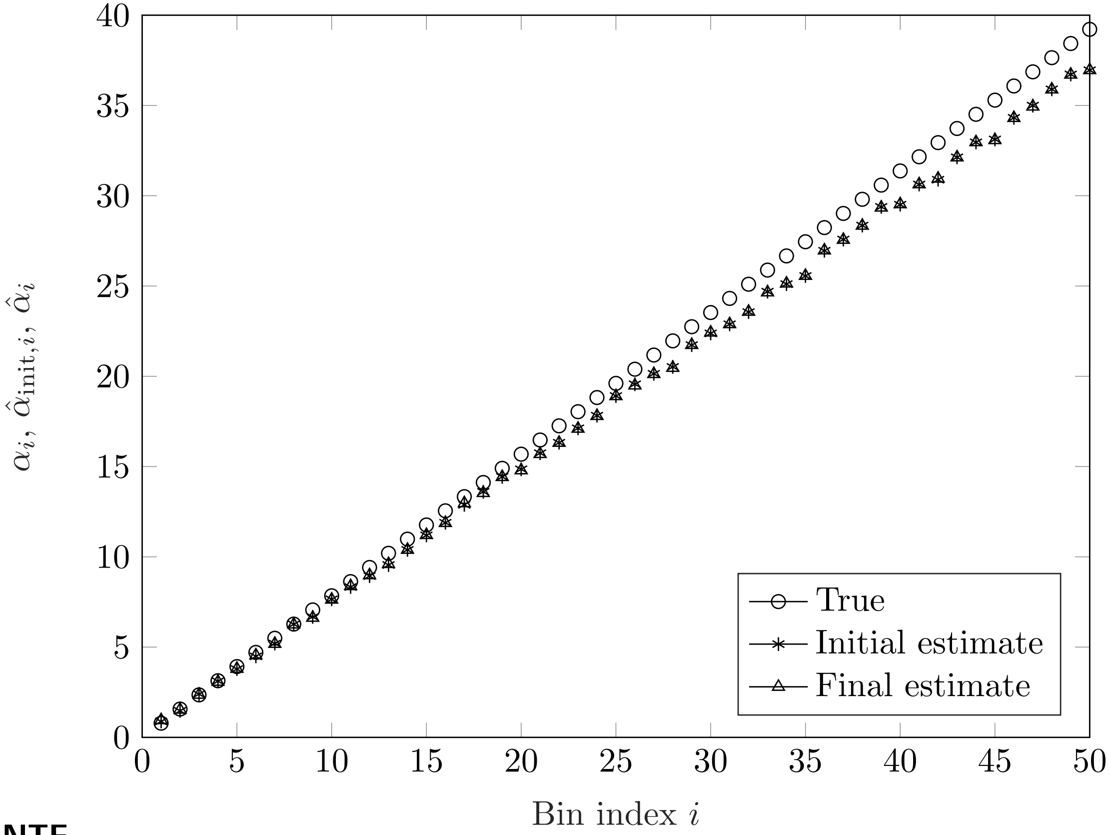
$$\hat{\boldsymbol{\alpha}} = \underset{\alpha}{\operatorname{argmax}} \sum_{j=1}^{J} \log(\mathrm{Dir}(\hat{\boldsymbol{p}}_j | \boldsymbol{\alpha}))$$

Number of users in training set.

Estimated bin probability vectors of individual users in training set. Computed from a collection of histograms per user.

Optimisation by Nelder-Mead uphill simplex method

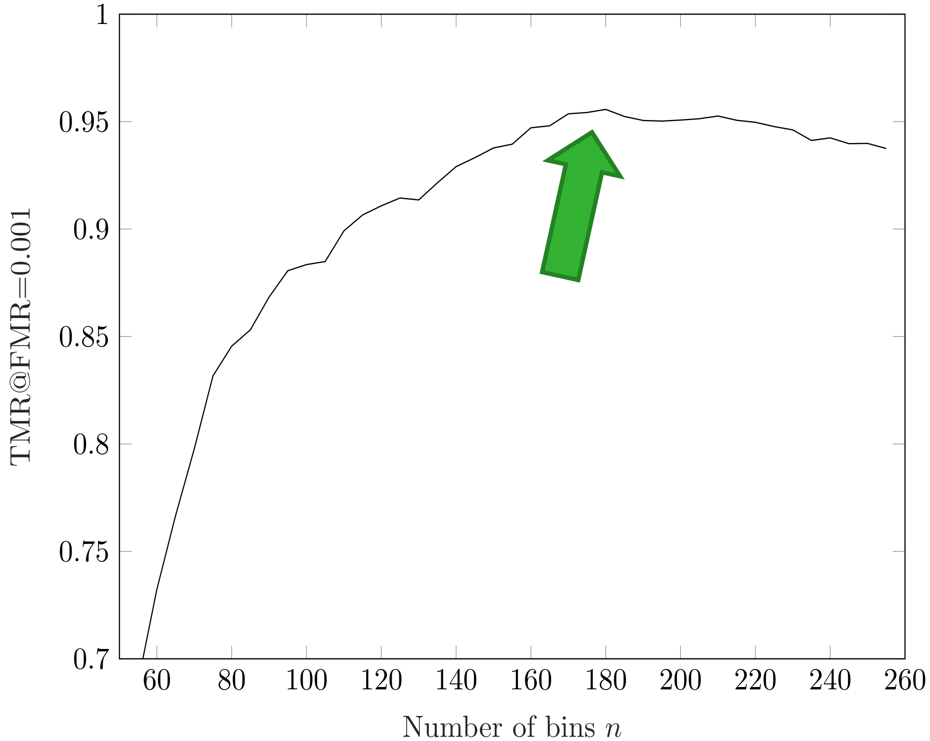
ML estimation of Dirichlet parameters

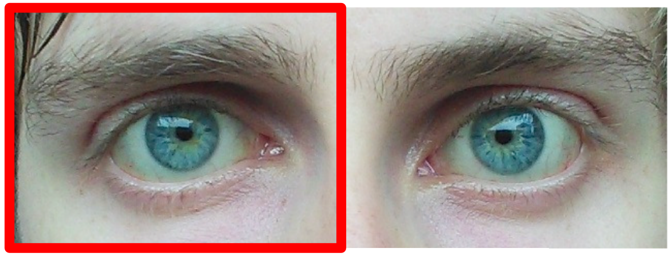


Feature selection

Periocular region

- Before training discard bins with higher probabilities
  - Improves generalisation properties
  - Reasons:
    - Bins with higher bin probabilities less discriminative?
    - Bins with higher bin probabilities restrict outcome space?
    - …

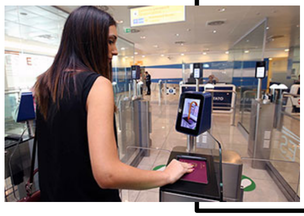



Left periocular region

Experiments

- 100 subjects for training, 46 for testing, BSIF features,15 histograms/subject
- Number of bins 256, reduced to 175

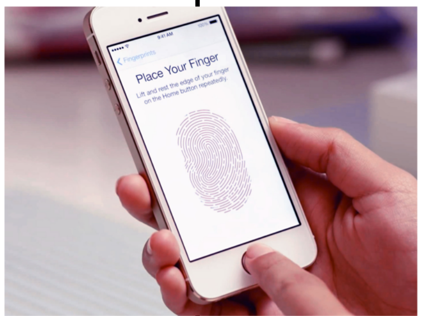
1:1 comparisons:



$$\text{Likelihood ratio} \quad : \quad \mathrm{lr}(\boldsymbol{x}, \boldsymbol{y} | \hat{\boldsymbol{\alpha}})$$

$$\text{Chi} - \text{square} \quad : \quad 2\sum_{i=1}^{n} \frac{(x_i - y_i)^2}{(x_i + y_i)}$$

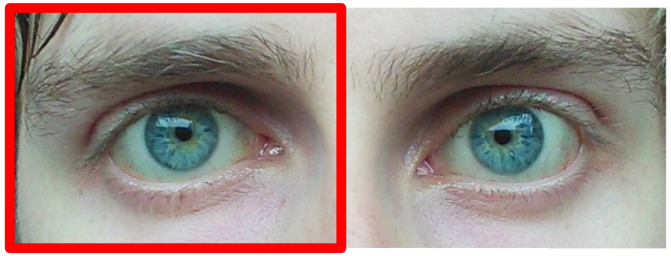
User-specific comparisons:



$$\text{User} - \text{specific likelihood ratio} \quad : \quad \mathrm{lr}(\boldsymbol{x} | \hat{\boldsymbol{q}})$$

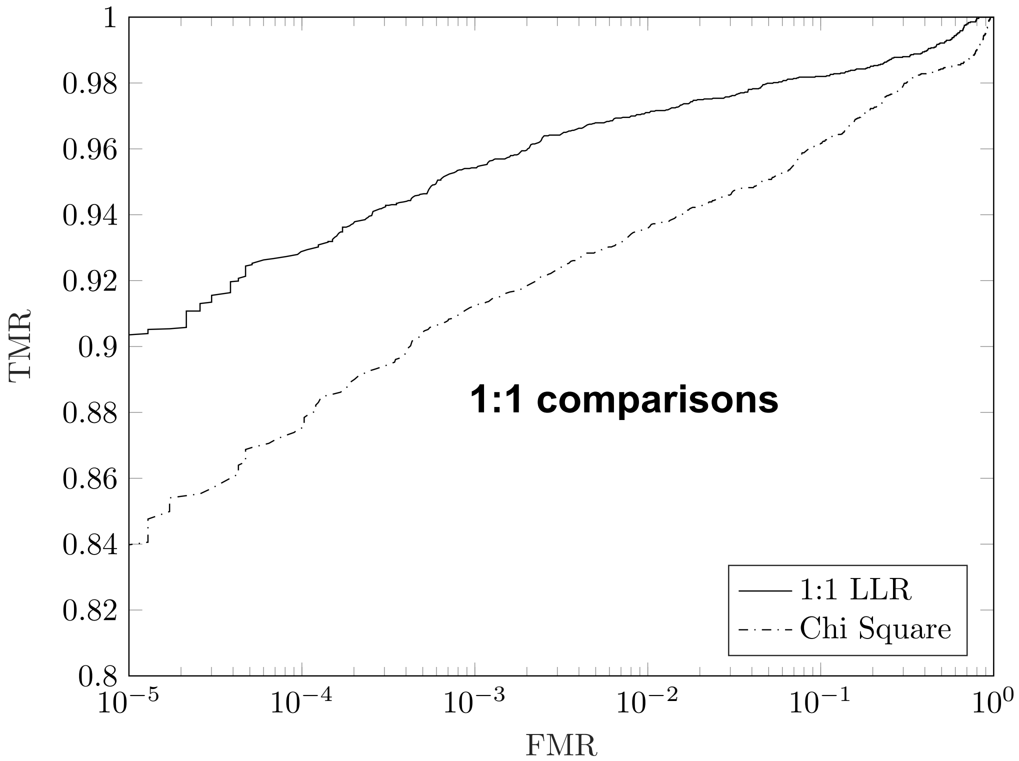
$$\text{Multiple} - \text{enrolment likelihood ratio} \quad : \quad \mathrm{lr}\left(\boldsymbol{x}, \sum \boldsymbol{y} | \hat{\boldsymbol{\alpha}}\right)$$

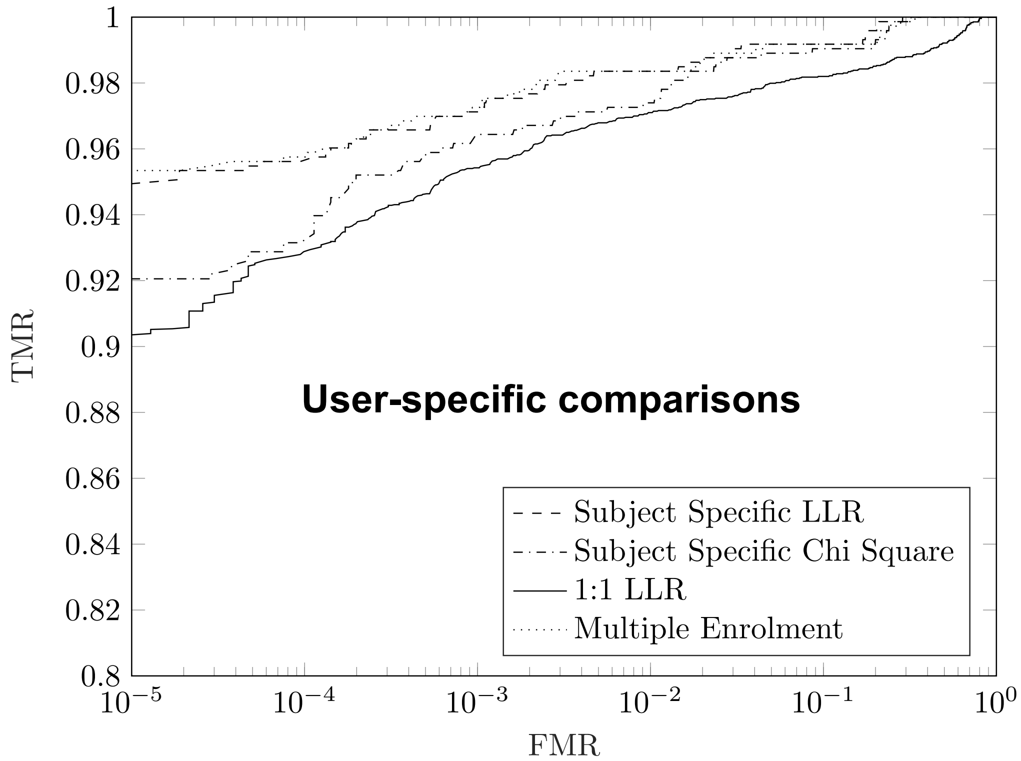
$$\text{Chi} - \text{square} \quad : \quad 2\sum_{i=1}^{n} \frac{(x_i - N\hat{q}_i)^2}{N\hat{q}i}$$



Left periocular region

Recognition performance





Conclusions

- New LR classifier for the comparisons of histogram features.
- ML estimation of background model, using Nelder-Mead optimisation.
- User-specific comparison outperforms 1:1 comparison
- Outperforms commonly used Chi-square classifier on periocular data.
  - Now testing on other datasets.
- Feature selection is beneficial but needs further research.

UNIVERSITY OF TWENTE.

Questions?