Face Recognition Across Ages

Bilgin Esme, Apr 2010

Face recognition across ages

Face recognition

Age related concerns
In face recognition

Face recognition
across ages

Non-generative approaches

Why non-generative approach?

- Generative approaches yield interesting contributions to various problems but have some serious problems in this specific problem

- People age very differently, so an aging simulation may estimate totally different future face, it's the nature of human growth.

- Some very abrupt changes seriously change the aging process. Such as sudden weight gain, depression, drug use etc.

- Generative approaches require an estimation of the 'target age', which we may have no idea at all.

- So bypassing the age simulation and trying to find some static facial patterns, that do not change across time is worth contemplating.

The Core Idea

- Obviously, the features do not drift independently
- Feature drift pattern on facial area has some distinct characteristics that doesn't change while aging
- Depending on the underlying shape and muscle structure of the individual, there's some coherency among these drifts

Following Biswas et.al.

BS.Biswas, G.Aggarwal, N.Ramanathan, R.Chellappa,

'A Non-generative approach for face recognition across aging', *2008*

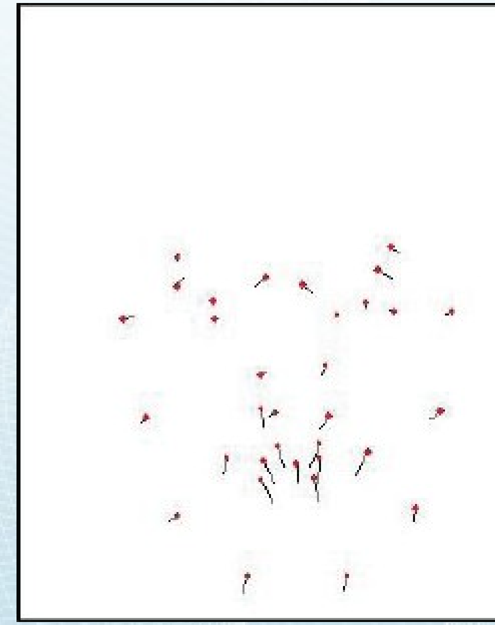
- **This is not a new idea**

- **The coherency of feature drifts are used in many researches including:**

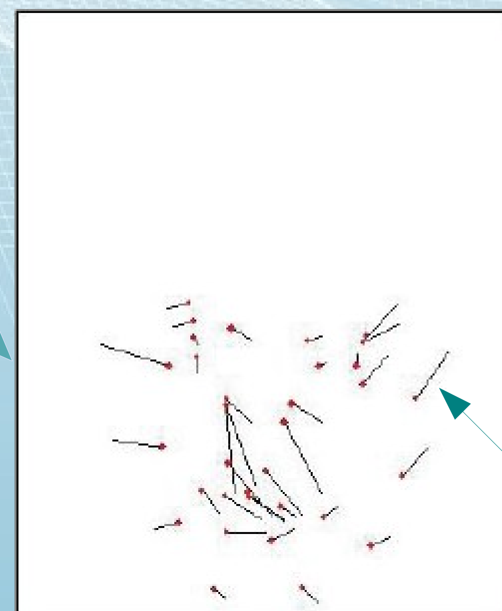
B.Li & R.Chellappa, 'Face verification through tracking facial features', *Journal of the Optical Soc. Of America, 2001*

Coherency of Feature Drifts

coherent



incoherent



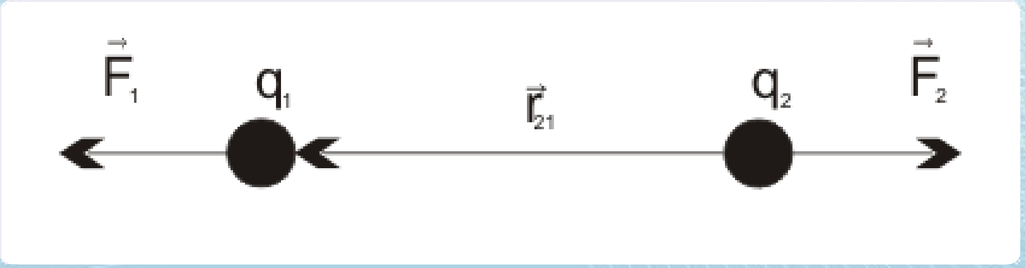
A single feature drift

Model for calculating FDs

- The model is adopted from the theory of electrostatics.

The potential energy between two charges q_i and q_i , separated by a distance r_{ij} is given by:

$$U_E = k_e \frac{q_i q_j}{r_{ij}}$$



k_e is the Coulomb constant.

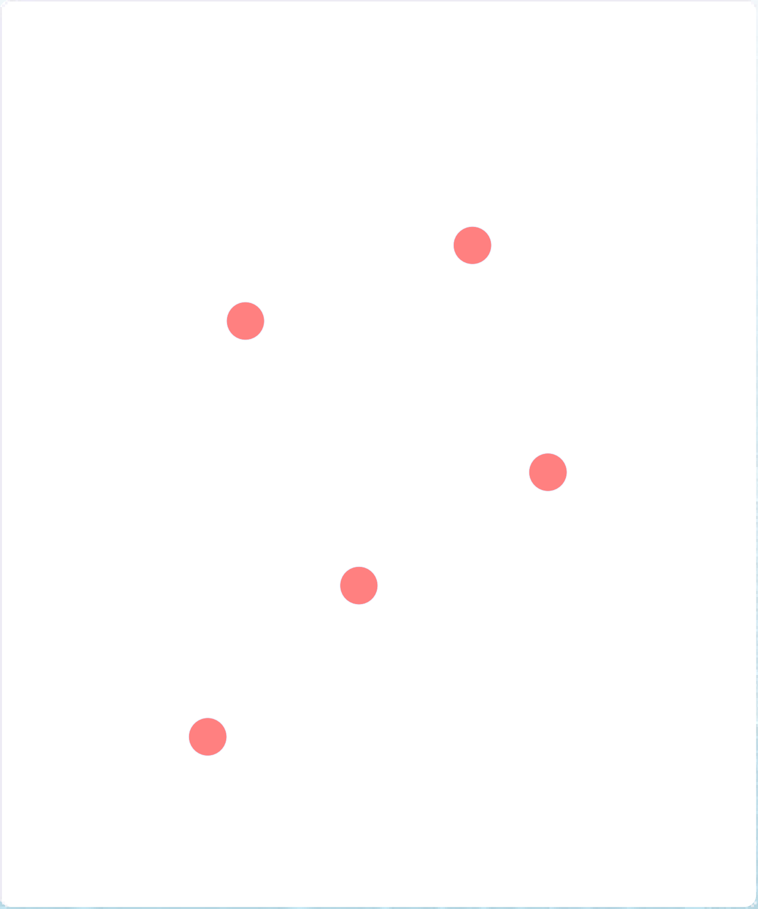
Model for calculating FDs

- The combined potential energy is calculated by using the superposition property of electrostatic charges.

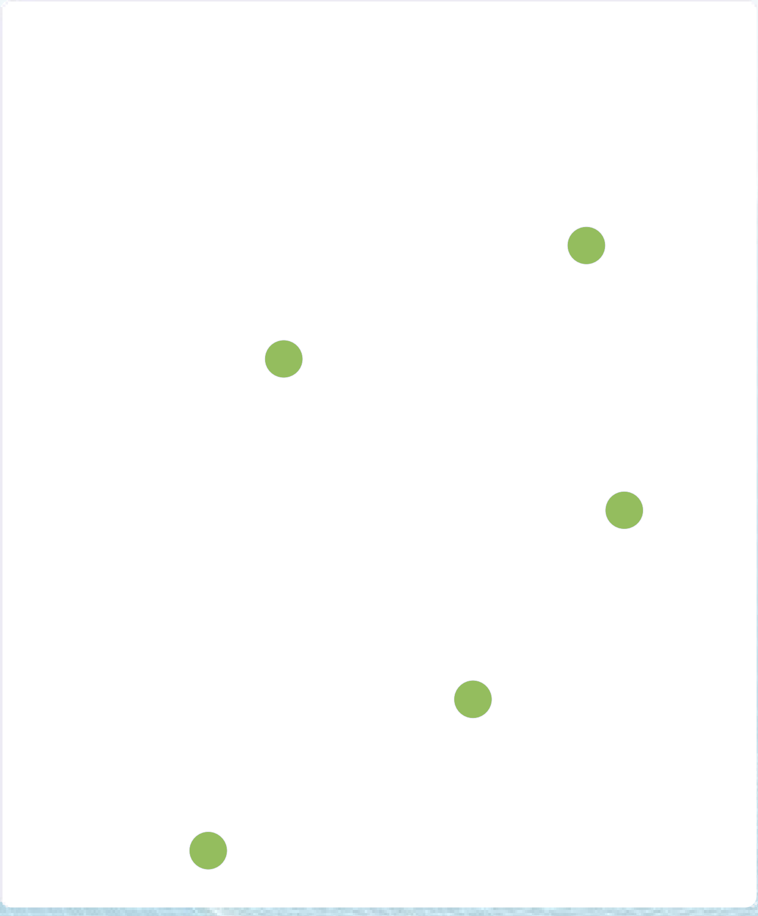
$$U_E = k_e\left(\frac{q_1 q_2}{r_{12}} + \frac{q_1 q_3}{r_{13}} + \frac{q_2 q_3}{r_{23}} + \ldots\right)$$

- The more stable system means the minimum combined potential energy.

Application of the model

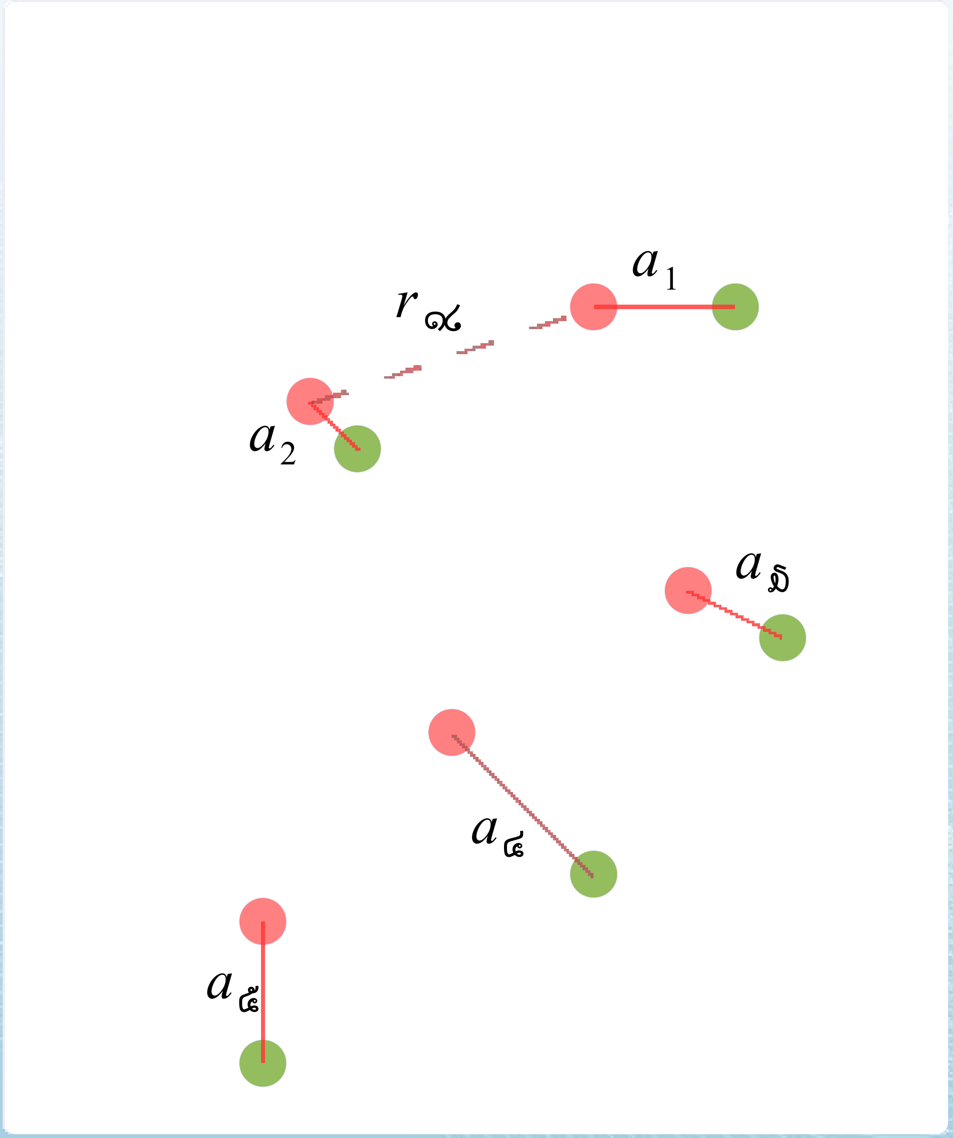


Feature Set 1



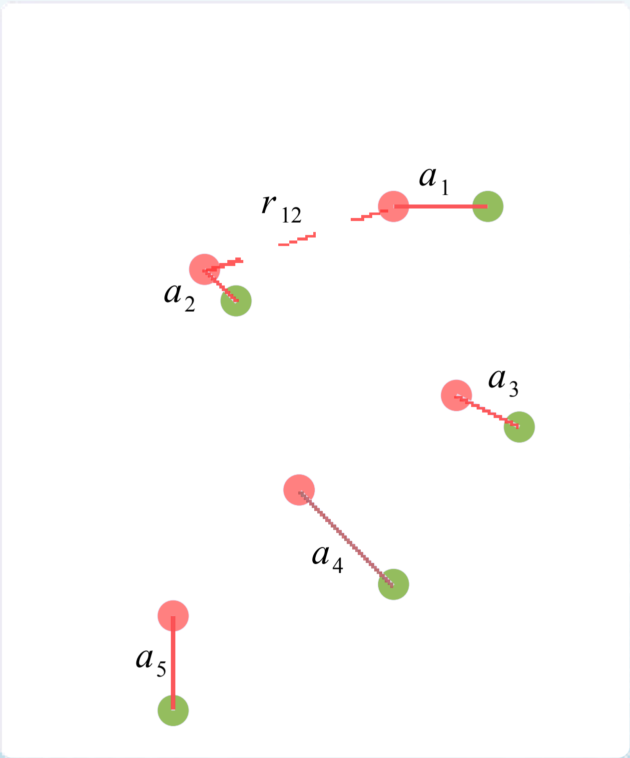
Feature Set 2

Application of the model



Feature Drift Map

Application of the model



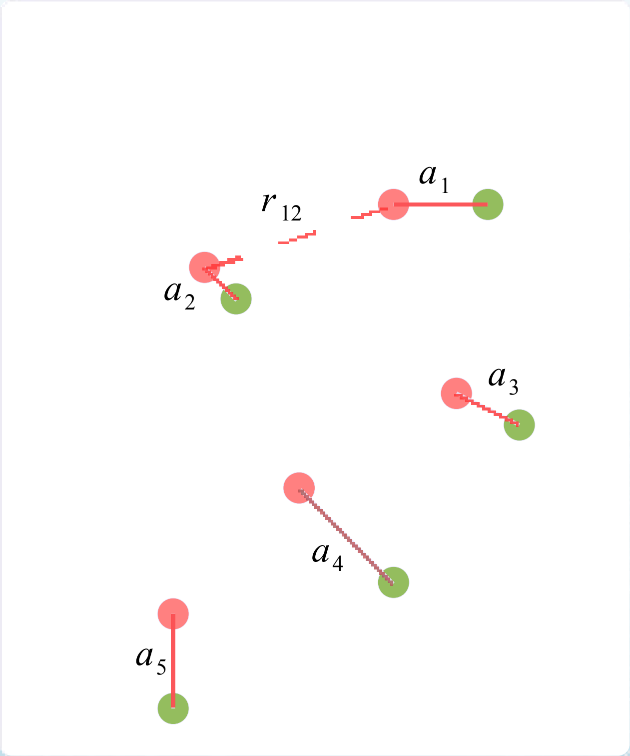
Feature Drift Map

- Based on the 'electrostatic charges' analogy, a measure of incoherency between tow feature drifts is defined as:

$$U_{ij} = \frac{|a_i - a_j|}{r_{ij}}$$

$|a_i - a_j|$ is the magnitude of the vector difference between two feature drifts.

Application of the model



Feature Drift Map

- Following the superposition property the combined potential energy of 'K' feature drifts is calculated as:

$$C = \sum_{i=1}^{K} \sum_{j=i+1}^{K} U_{ij}$$

The lower this figure is, it's most probable that two images belong to the same person.

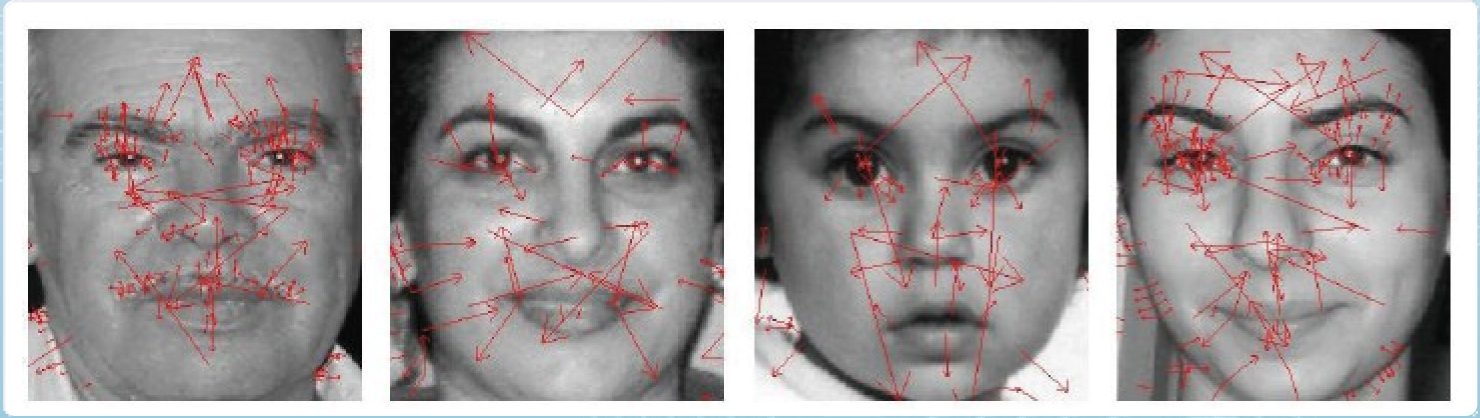
Obtaining the feature drifts

- **Using SIFT features**

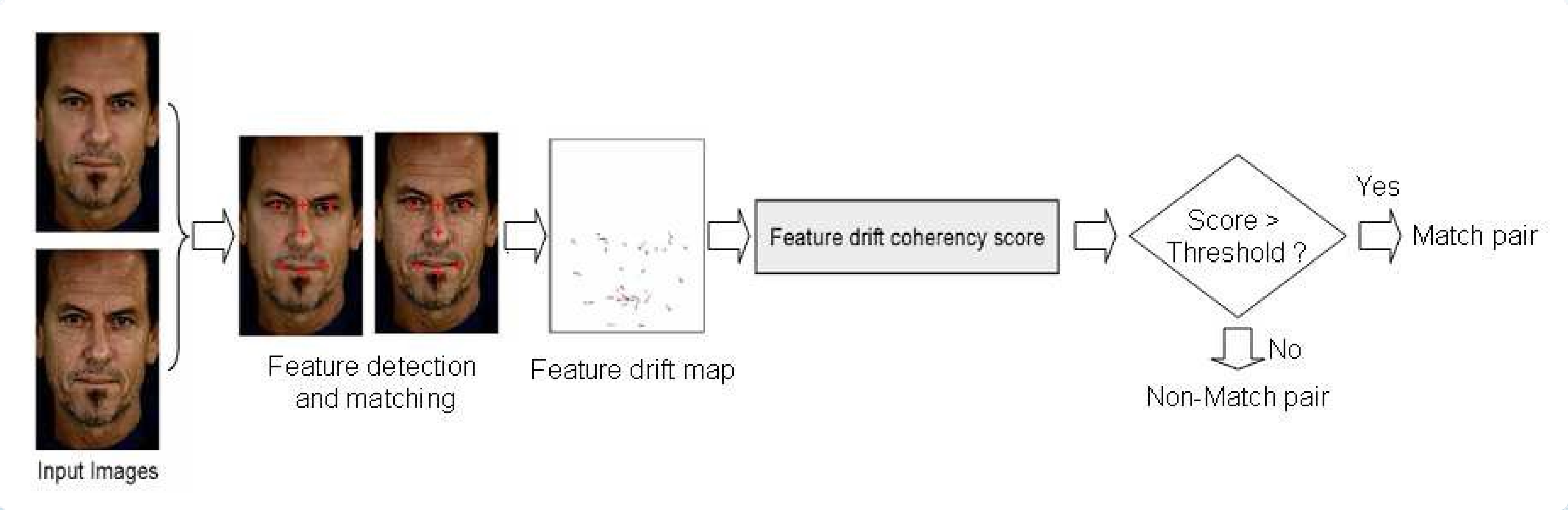
SIFT Features
Scale-invariant feature transform

D.G.Lowe, 'Distictive image features from scale-invariant keypoints, 2004

- **Features are extrema points in scale-space of the image**
- **Each feature is characterized by 128 dimensional vector (gradient distribution around the feature)**



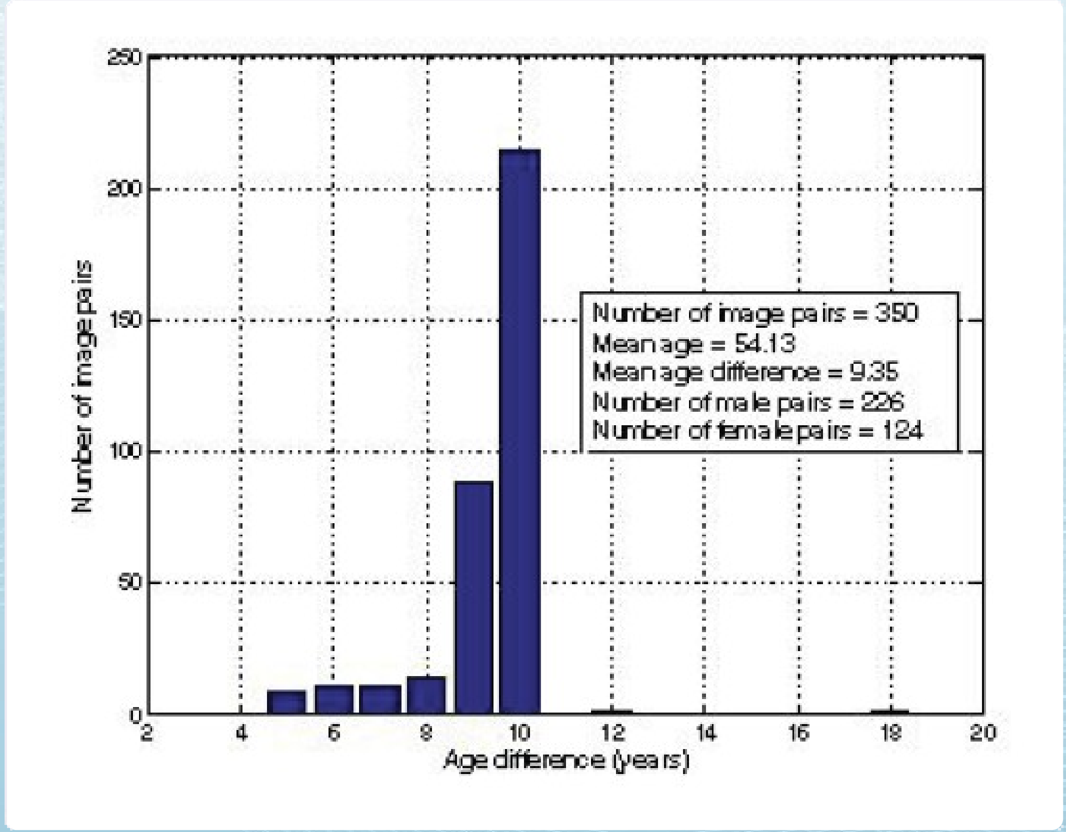
The algorithm



- A private passport database is used
- 350 pairs of images with age separation of over 9 years
- A ROC curve is obtained for a threshold range.

The algorithm

- A private passport database is used
- 350 pairs of images with age separation of over 9 years



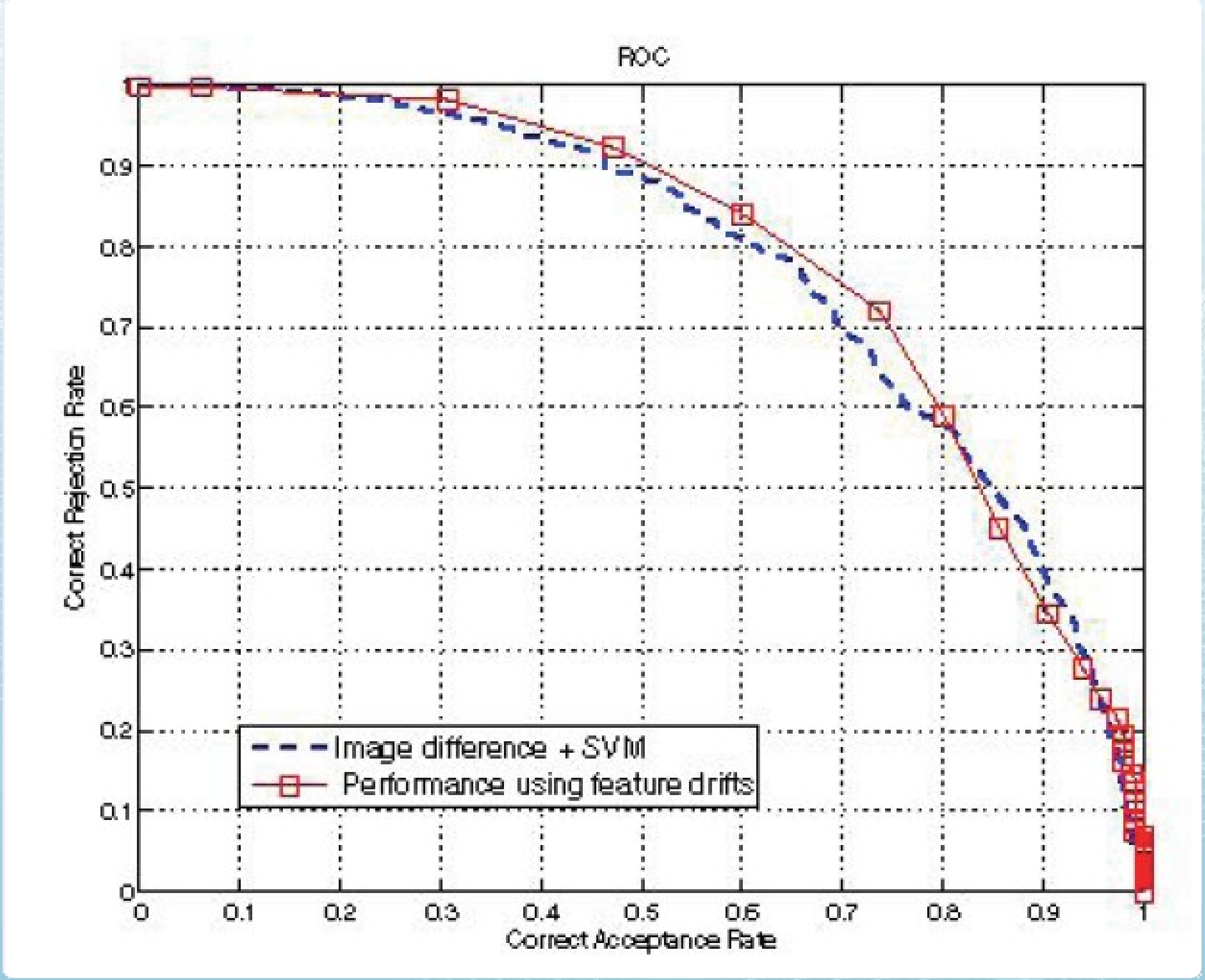
Experiment results

- **The algorithm is compared with a counterpart 'SVM+diff' algorithm**

SVM+diff

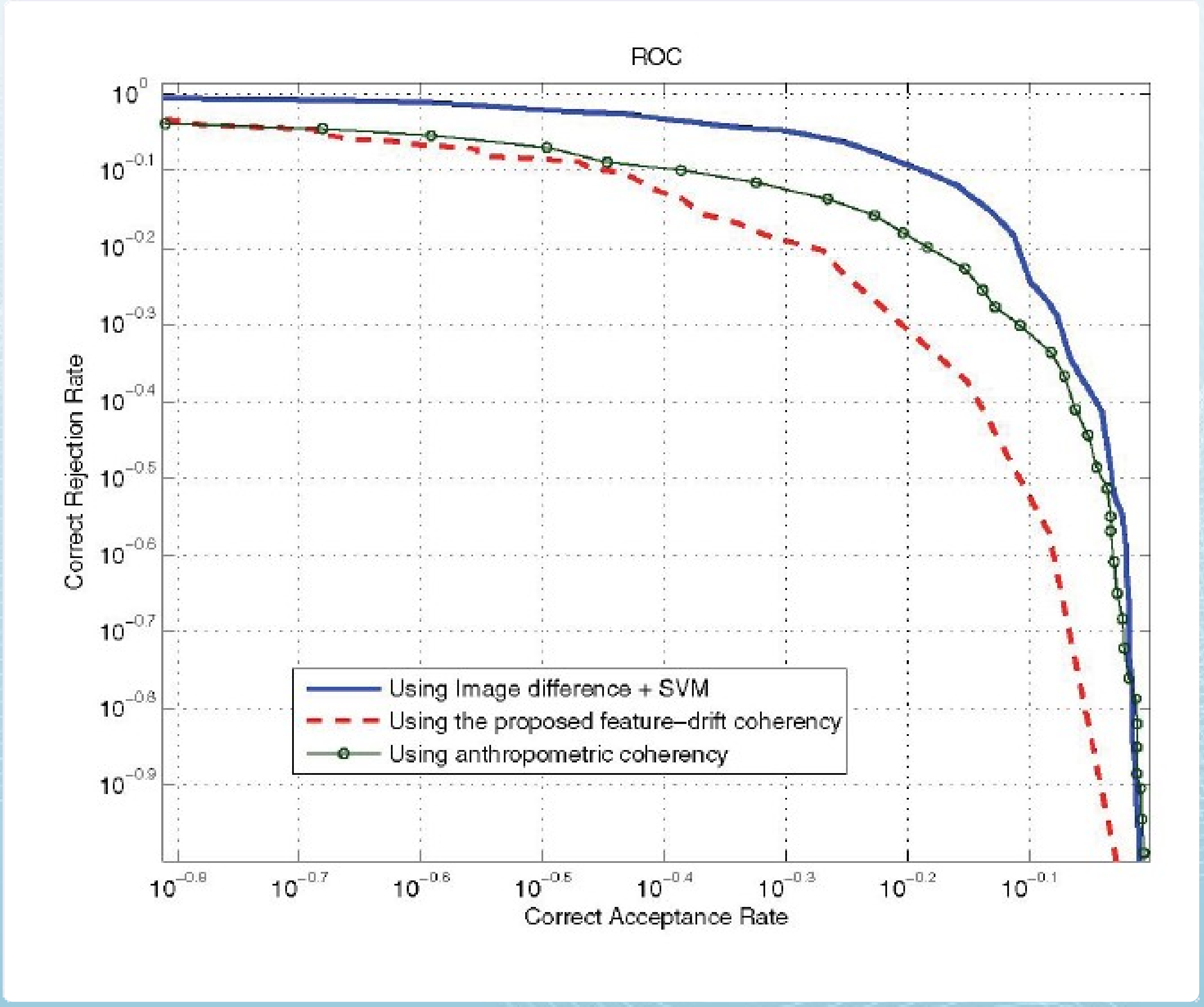
P.J.Phillips, 'Support Vector Machines applied to face recognition', 1999

- Uses differences of normalized images
- An elliptic region is cut and re-sized to 80x70
- Normalized to have zero mean and unit variance



Experiment results ???

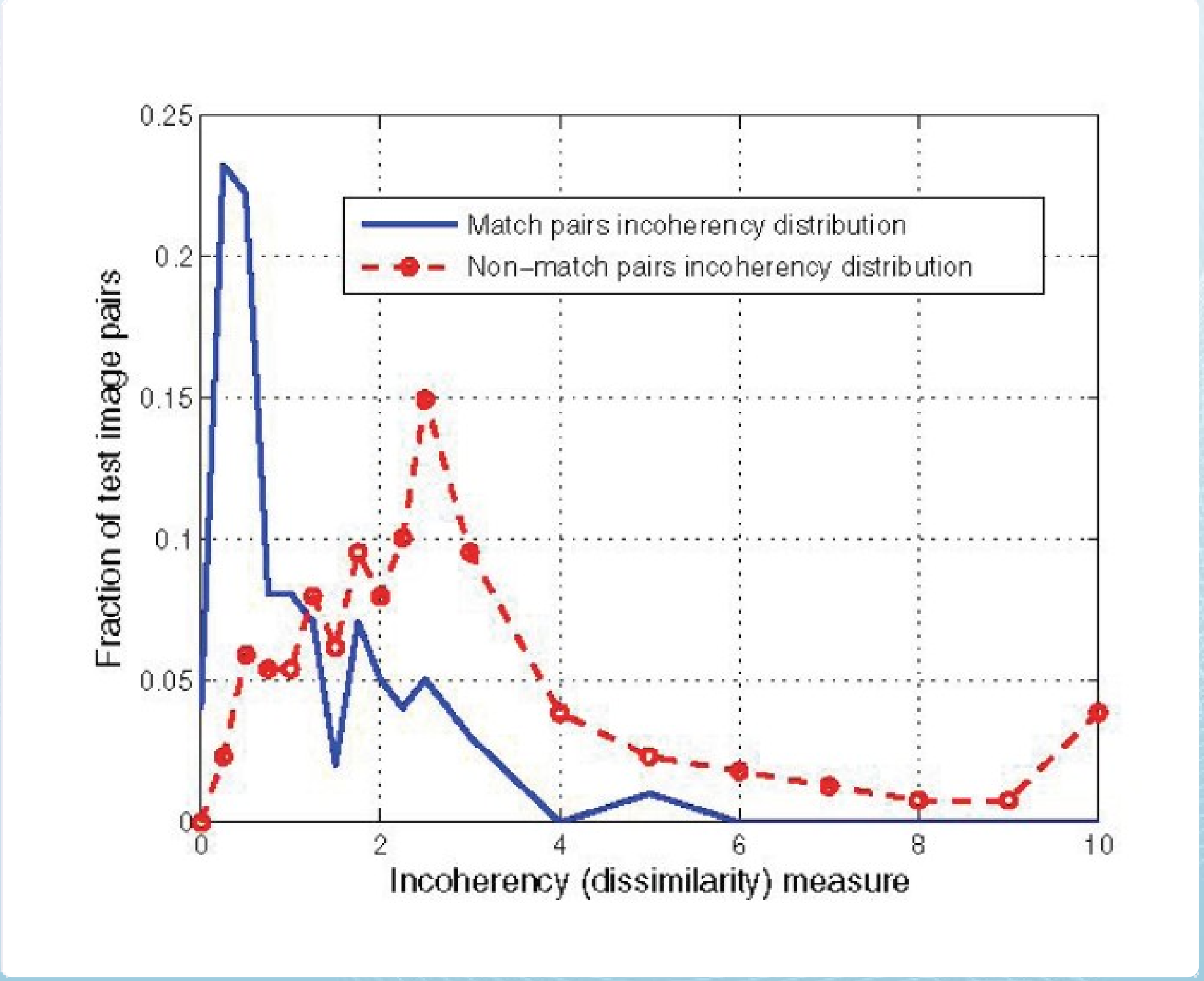
- The results for child images



Experiment results

• Genuine and impostor incoherency score distributions in the experiment

• It's clear that there's a meaningful separation between these distributions.



Discussions on results

- This specific algorithm proves that, non-generative approaches hold promising ground on face recognition across ages.

- The algorithm does well though it's very simple and only 10 features are used.

- This idea may be extended in many ways either to yield better results and/or pointing to different sub-problems

Conclusions & Future Work

- Using facial feature drifts for obtaining age-invariant signatures on human face is a promising research area.

- The 'coherency measure' is on its own a fertile area to grow different ideas.

- Modeling the coherency due to electrostatic theory is just a solution among many possible ones.

Future Work Ideas

- In this research, the features proposed by SIFT algorithm is used. These are not necessarily the best age-proof features. We can try some other features and take out the most appropriate ones.

- In this work, all feature drifts are regarded same. It may be more appropriate to use weights for each feature.

- The method used for penalizing the missing features seems primitive and arbitrary. By combining a weight assigning method, we can build a more robust scheme.

- Use 'substitute features' for missing ones. Determining these substitute also demands a some hard work.

Future Work Ideas

- Instead of calculating feature drifts, we can calculate the drifts of 'feature groups'. This may be useful, because 'sagging of a muscle' is generally a drift of multiple points.

- Apply the algorithm partially, such as on chin, cheeks, eye area; and then combine/compare results; and then conclude.

- What's the most critical/challenging age group? Find it and develop some additional measures for this age group.

- What happens when the age difference between two images is much bigger? e.g. > 20 years