The Empathic Visualisation Algorithm (EVA): Chernoff faces revisited

Andreas Loizides
(Mel Slater)

P.A. College
University College London

Motivation

Balance Sheet					
Capital & Reserves					
ORDINARY SHARE CAPITAL	131000	131000	132000	141000	141000
SHARE PREMIUM A/C	836000	840000	856000	1441000	1460000
OTHER RESERVES	39000	104000	163000	230000	308000
PROFIT & LOSS A/C	2380000	2629000	2690000	2381000	2644000
EQUITY CAP. AND RESERVES	3386000	3704000	3841000	4193000	4553000
PREFERENCE CAPITAL	0	0	100000	200000	325000
TOT. SHARE CAPITAL &RESERVES	3386000	3704000	3941000	4393000	4878000
Fixed Assets					
INTAGIBLE					
TANGIBLE	509000	534000	585000	715000	698000
INVESTMENTS	25312000	32374000	35298000	39782000	46558000
OTHER	1415000	1433000	2074000	2038000	2908000
	27236000	34341000	37957000	42535000	50164000
Current Assets					
STOCKS	0	0	0	26000	14000
DEBTORS	50857992	53668000	57512992	71361992	86548992
INVESTMENTS	0	0	0	0	0
OTHER	1080000	1620000	1753000	1880000	2023000
CASH	757000	598000	391000	339000	1957000
	52694992	55886000	59656992	73606992	90542992
Current Liabilities					
PROVISION FOR TAX	442000	498000	496000	299000	266000
PROVISION FOR DIVIDENDS	129000	158000	191000	245000	290000
CREDITORS <1 YEAR	55757992	60526000	64129992	72659000	91462992
OTHER	11373000	11372000	15334000	22796000	25594000
	67701992	72554000	80150992	95999000	1.18E+08
Net Current Assets	-15007000	-16668000	-20494000	-22392008	-27070000
Total Asset Less Current Liabilities	12229000	17673000	17463000	20142992	23094000
Long Term Liabilities					
PROVISIONS	296000	399000	630000	970000	1144000
LOAN CAPITAL	8547000	13570000	12892000	14780000	17072000
OTHER					
	8843000	13969000	13522000	15750000	18216000
	3386000	3704000	3941000	4392992	4878000
Profit & Loss					
OPERATING PROFIT-ADJ	-446000	-504000	-1169000	-611000	-394000
TOTAL NON-OPERATING INCOME	1247000	1514000	2337000	2019000	2011000
TOTAL INTEREST CHARGES	66000	80000	134000	167000	193000
PROFIT BEFORE TAX	735000	930000	1034000	1241000	1424000
TAX	310000	319000	337000	403000	480000
PROFIT AFTER TAX	425000	611000	697000	838000	944000
ORDINARY DIVIDENDS	184000	233000	288000	360000	434000
TO SHAREHOLDERS FUNDS	241000	378000	409000	478000	510000

- For 2,3 variables graphs and spreadsheets are very effective
- What about when you have 5, 10, 20, 70 variables? The number of variables is too large to be directly encoded into an orthogonal visual structure

UCL

COLLEGE

Motivation: Problems with existing techniques

- Complexity of systems increases with dimensionality
- Learning time is required from the users in order to understand and be able to interact with these mediums effectively
- Very hard to visualise relationships beyond quadratic
- Don’t get a meaningful holistic view of your data

UCL

COLLEGE

Chernoff Faces

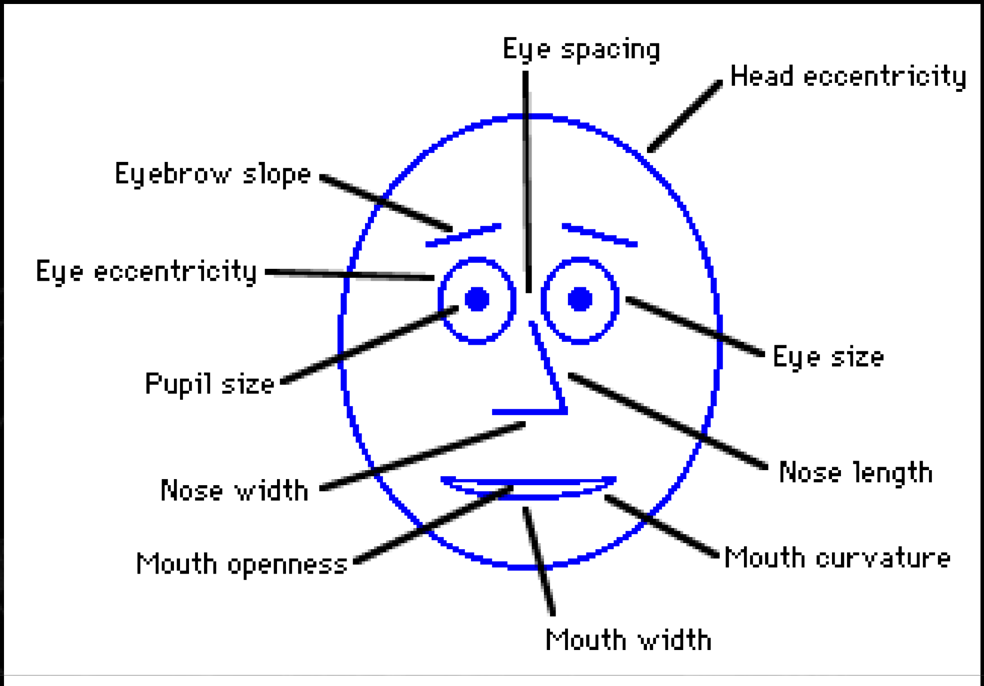
- Mapping data variables into features of schematic faces.

His idea capitalises on 2 important principles:

- Familiarity with human faces and our ability to process even the smallest changes due to everyday interaction.
- A face evokes an emotional response in us

- Psychological research indicates that the human face can be an excellent abstraction of data [Walker, Wilkinson, Homa]

Chernoff Faces



Problems:

- Variables treated uniformly
- User training is required
- You don't see the actual values
- Loses effectiveness for extreme values
- Subjectiveness of visual structure

UCL

COLLEGE

Empathic Visualisation Algorithm

- Differs from Chernoff faces
  - Automatic derivation from data
  - Emotions in the face reflect importance in the data
- Faces are only one example

Empathic Visualisation Algorithm: EVA

- $n \times k$ data matrix X of n cases on k "quantifiable" variables $x_1, x_2, \ldots, x_k$

$$\underset{(nxk)}{Matrix}$$

$$\begin{array}{cc} & \text{Variables} \\ \text{cases} & \begin{bmatrix} X_{11} & X_{12} & X_{13} & \cdots & X_{1k} \\ X_{21} & X_{22} & \cdots & \cdots & X_{2k} \\ \vdots & \vdots & & & \vdots \\ X_{n1} & X_{n2} & \cdots & \cdots & X_{nk} \end{bmatrix} \end{array}$$

Objective

Overall: Construct a visualisation such as the salient features of the data can intuitively be recognised by an observer and representation gives an overall view of the data set.

Two further objectives:

- Naturalistic visual representation; something encountered in everyday life
- Automatic Mapping; semantically important features in the data are mapped to perceptually or emotionally important features of the visual structure -> Visual Homomorphism

$$\begin{bmatrix} X_{11} & X_{12} & X_{13} & \cdots & X_{1k} \\ X_{21} & X_{22} & \cdots & \cdots & X_{2k} \\ \vdots & \vdots & & & \vdots \\ X_{n1} & X_{n2} & \cdots & \cdots & X_{nk} \end{bmatrix}$$

$$(X)$$

Value system (chosen by user)	Given
Profit over Total Assets	$v_1(X)$
Current Ratio (short term liquidity)	$v_2(X)$

$$(\Omega)$$

Global characteristics of Visual Structure	Measured
I am happy	$e_1(\Omega)$
I am angry	$e_2(\Omega)$

UCL

COLLEGE

Visual Homomorphism

Find a mapping:

$$\mu : X \rightarrow \Omega$$

- Global characteristics of the data correspond to emotional expressions of the face:
  - The company is profitable – I am happy
  - Short term liquidity is a problem – I am angry

How?

DATA

Feature functions

$$f_1(X) = x_7 x_3 e^{2x_9}$$

$$f_2(X) = 10x_1x_4 + \ln(7.3x_8)$$

....

FACE

muscle contractions

Zygomatic major $\varphi_1(\Omega)$

Left frontalis inner $\varphi_2(\Omega)$

....

- Feature functions determine muscle contractions
- One set of r feature functions $\rightarrow$ face
- Geoface 2

UCL

COLLEGE

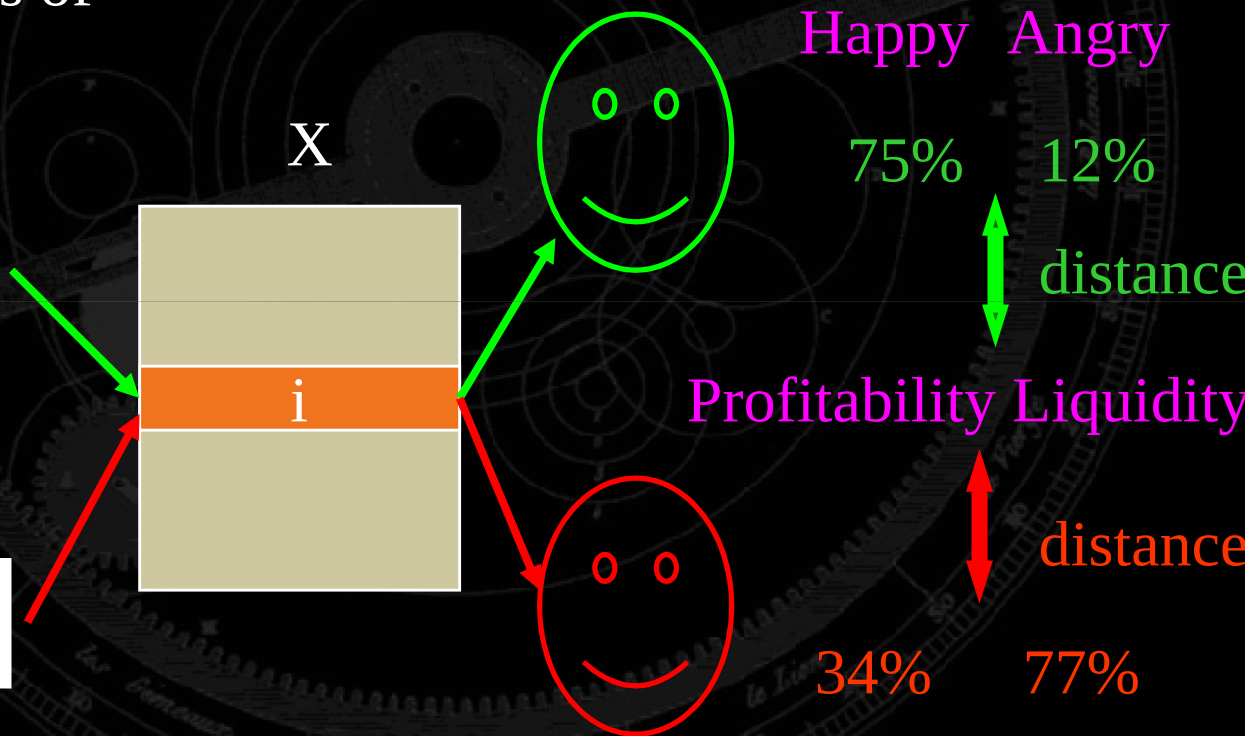
How do we choose the feature functions?

Population of N sets of
Feature functions

$$f_{11}, f_{12}, \ldots, f_{1r}$$

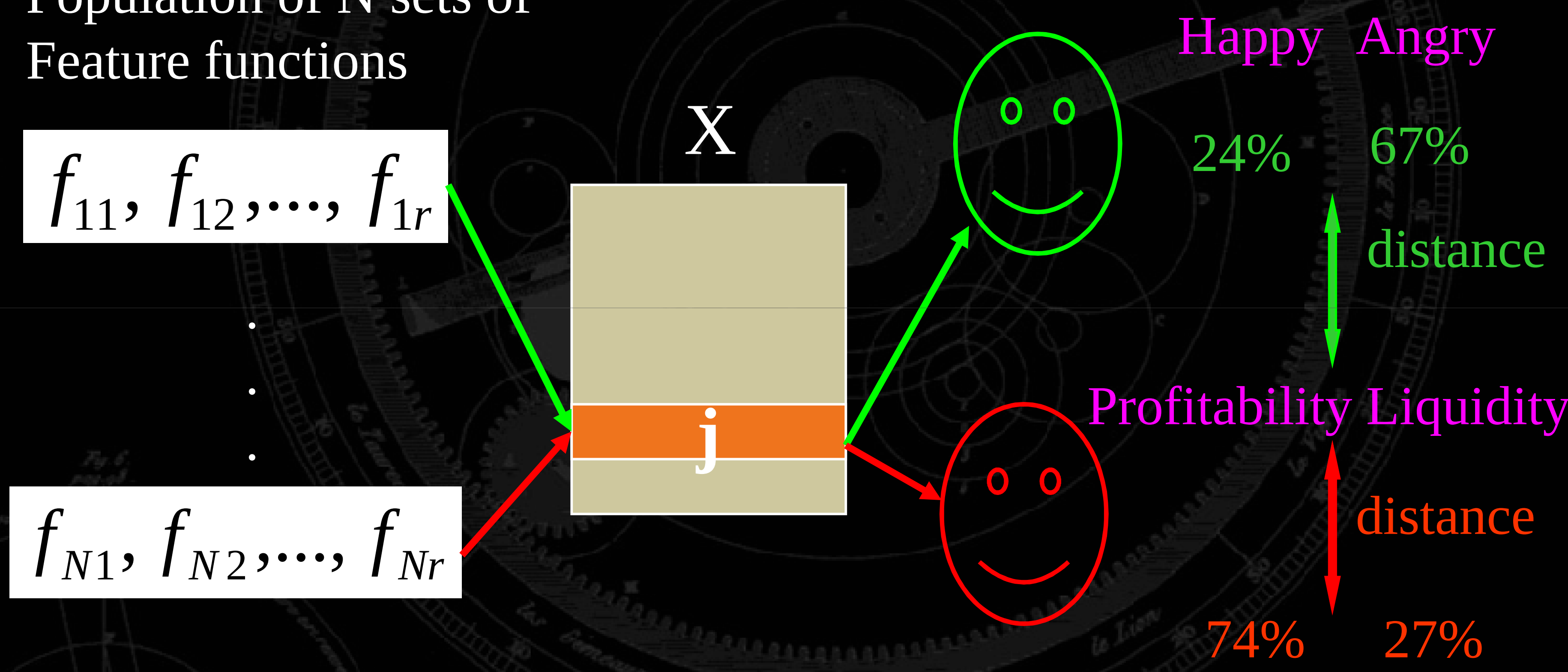
.
.
.

$$f_{N1}, f_{N2}, \ldots, f_{Nr}$$



Same for row j

Population of N sets of Feature functions



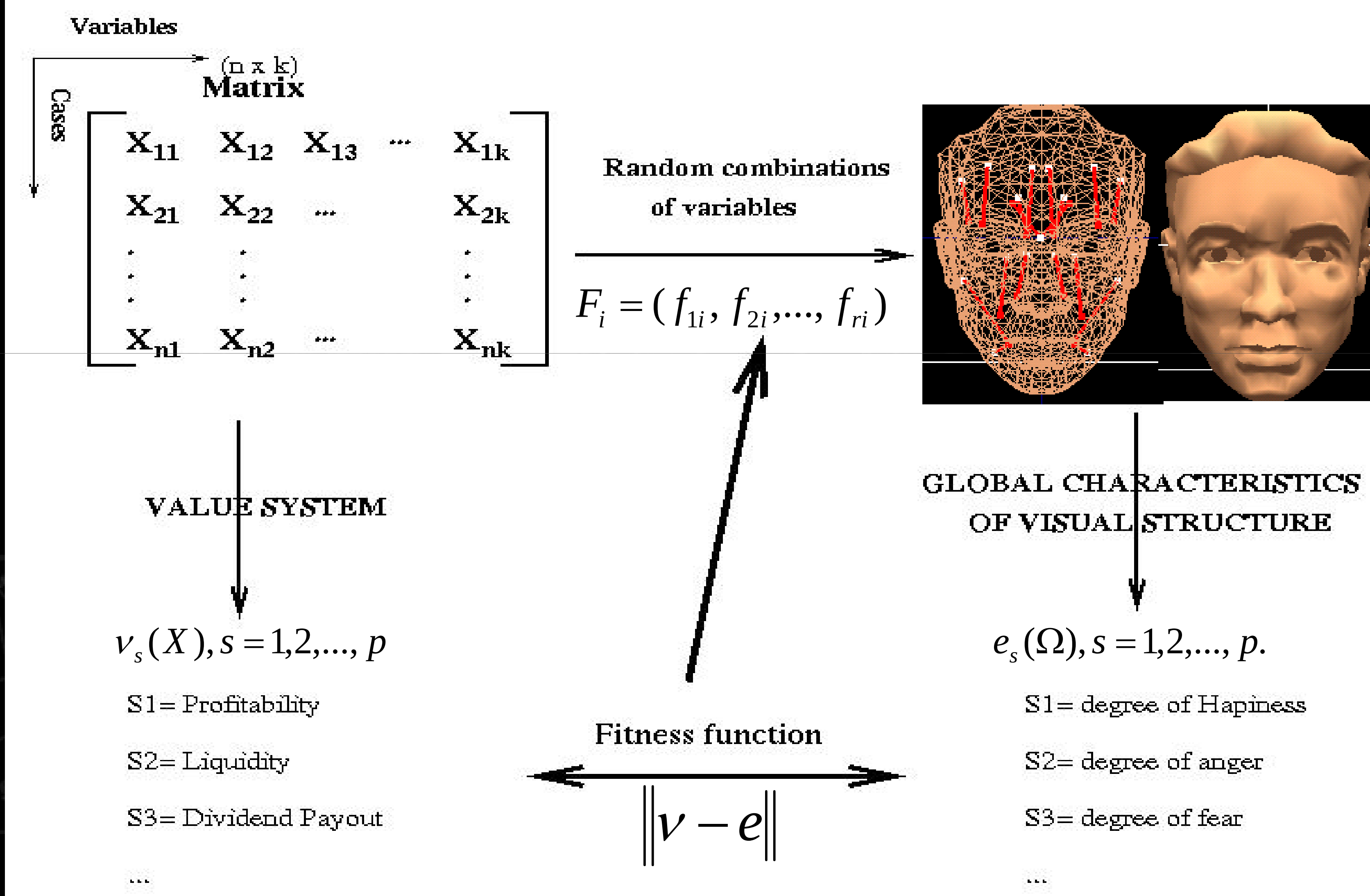
Error Measurement

- Error
  - for an individual set of feature functions = sum of squared distances over all rows
- Fitness
  - For an individual set of feature functions may be derived from the errors
  - Used to determine probabilities of selection

Genetic Program

- Choose first generation of feature functions as random functions over the variables
- Next generation uses probabilities based on fitness, determines survival and
  - Selection for reproduction and Mating
- Repeat for each new generation
- Greater fitness $\Rightarrow$ greater match between global characteristics of data and characteristics of face

From Data to Visual Structure: An automatic mapping



FOPS – Fear of Public Speaking Experiment

- 40 users gave talk in a VR environment under different settings.

Condition	Empty		Audience	
	M	F	M	F
Phobic	5	5	5	5
Confident	5	5	5	5

UCL

COLLEGE

13 variables used + heart rate

1. Gender (0/1)
2. Anxiety Level (C/P)
3. FNE (out of 30)
4. STAI (20-80)
5. PRCS (out of 30)
6. Audience Type (0/1)
7. Presentations (0/1)
8. Subject (1-7)
9. Prepared (1-7)
10. Emotions before *
11. Emotions after (0-100)
12. Self Rate (0-100)
13. Somatic

* Drawn out of six scales of adjectives (happy, sad, fear etc.)

UCL

COLLEGE

QUESTION:

Does the statistical analysis on the gathered data match the emotional aspects of the faces produced by the use of EVA?

Objectives:

- Test validity of EVA when compared to statistical analysis.
- Test the method in a different context

Data Value System vs Visual Structure Characteristics

Value System (for each user given a 0-100 value)

Global Characteristics (0-100 for each user)

- Somatic (v1) ⟷ Degree of Happiness (e1)
- MPRCS (v2) ⟷ Degree of Fear (e2)
- (Somatic x MPRCS) – to allow for the interaction effect (v3) ⟷ Degree of Anxiety (e3)

$$\sum_{d=1}^{40} \sum_{s=1}^{3} (v_{ds} - e_{ds})^2$$

UCL

COLLEGE

RESULTS

FOPS Results

Conclusions

- Automatically map multidimensional data to human faces using GP
- You get a quick understanding of the overall impact of the data
- Complementary not an alternative to statistical analysis and other techniques
- Visualising the data matrix as a whole

- QUESTIONS?