Representation and Recognition in Vision

Shimon Edelman
School of Cognitive and Computing Sciences
University of Sussex
Falmer, Brighton BN1 9QH
UK

email: shimone@cogs.susx.ac.uk

May 12, 1998

Contents

1 The problem of representation **19**

1.1 A vision of representation 19

1.2 Reconstruction 21

1.3 Representation without reconstruction 24

1.3.1 The feature detector redux 24

1.3.2 The challenge 25

2 Theories of representation and object recognition **29**

2.1 Recognition-related tasks that require representation 29

2.1.1 Identification and generalization 31

2.1.2 Categorization 31

2.1.3 Analogy 34

2.2 A formalization of the notion of representation 36

2.2.1 The Problem of Representation 36

2.2.2 Representation as a mapping 37

2.2.3 First and second-order isomorphism 37

2.3 Computational theories of recognition 42

2.3.1 Reconstructionist theories: a brief historical perspective 42

2.3.2 Structural decomposition theories 43

2.3.3 Theories based on geometric constraints 47

2.3.4 Multidimensional feature spaces 49

3 $\mathcal{S}$-isomorphism: The theory **57**

3.1 Similarity as proximity in a metric space 58

3.1.1 Some common objections 59

3.1.2 A metric similarity space as a working hypothesis 61

3.2 Shape spaces 62

3.2.1 Kendall's shape space 62

3.2.2 Transformations and deformations . . . 63
3.2.3 Best-correspondence distance . . . 64
3.2.4 An objective shape space . . . 65
3.3 Parameterization of distal shape space . . . 66
3.3.1 Scope of parameterization . . . 67
3.3.2 Dimensionality of parameterization . . . 69
3.4 Distal to proximal mapping . . . 71
3.4.1 Levels of representation of similarity . . . 73
3.4.2 The components of the mapping F . . . 74
3.4.3 Constraints on F . . . 76
3.4.4 Implications . . . 78

4 $\mathcal{S}$-isomorphism: An implementation 83
4.1 Task-dependent treatment of the measurement space . . . 85
4.1.1 Identification ("is this an image of object X?") . . . 85
4.1.2 Recognition ("is this an image of something I know?") . . . 92
4.1.3 Categorization ("what is this thing?") . . . 93
4.2 Categorization as navigation in shape space . . . 100
4.2.1 Defining the frame of reference . . . 100
4.2.2 Active landmarks . . . 101
4.3 Tuned units as active landmarks . . . 102
4.3.1 Relevance . . . 102
4.3.2 Coverage . . . 103
4.3.3 Smoothness . . . 103
4.3.4 Invariance to irrelevant dimensions . . . 103
4.3.5 Monotonicity . . . 105
4.3.6 Learnability . . . 106
4.4 The Chorus of Prototypes . . . 108
4.4.1 Persistent and ephemeral representations . . . 109
4.4.2 The underlying principles . . . 109
4.4.3 The applications . . . 110

5 $\mathcal{S}$-isomorphism: experiments 113
5.1 A Chorus of 10 reference-object modules . . . 114
5.1.1 The RBF module . . . 114
5.1.2 A multiple-module network . . . 116
5.2 Experimental results: recognition-related tasks . . . 118
5.2.1 Identification of novel views of familiar objects . . . 118

5.2.2 Categorization of novel object views . . . 119
5.2.3 Discrimination among views of novel objects . . . 124
5.3 Experimental results: analogy-related tasks . . . 126
5.3.1 Mechanism for supporting analogy . . . 130
5.3.2 Local viewpoint invariance for novel objects . . . 130
5.3.3 Recovery of a standard view for novel objects . . . 132
5.3.4 Recovery of pose for novel objects . . . 134
5.3.5 Prediction of view for novel objects . . . 134
5.3.6 Comparison with related methods . . . 134
5.4 Interim summary: Second-order Isomorphism and Chorus (SiC) . . . 137

6 Biological evidence **141**
6.1 Neurobiology: computing with receptive fields . . . 142
6.2 Neurobiology: low-level RFs . . . 143
6.2.1 Broad tuning . . . 143
6.2.2 Graded response . . . 143
6.3 Neurobiology: high-level RFs . . . 143
6.3.1 Selectivity to objects . . . 144
6.3.2 Ensemble encoding . . . 147
6.3.3 Selective invariance . . . 151
6.3.4 Plasticity and learning . . . 154
6.3.5 Speed of processing . . . 154
6.3.6 A summary of neurobiological support for SiC . . . 155
6.4 Psychophysics: Effects of viewpoint . . . 156
6.4.1 Canonical views . . . 156
6.4.2 Mental rotation . . . 158
6.4.3 Nonlinearities and the effects of practice . . . 161
6.4.4 A model of canonical views and mental rotation . . . 165
6.4.5 Viewpoint invariance as a tool for testing theories of representation . . . 167
6.5 Psychophysics: Veridicality of shape representation . . . 181
6.5.1 Perceived similarity as proximity in an internal metric space . . . 184
6.5.2 Veridical perception of distal shape contrasts . . . 185
6.6 Psychophysics: Effects of translation . . . 189
6.6.1 Translation invariance for line drawings of common objects . . . 190
6.6.2 Controlled synthetic animal-like shapes . . . 191

7 Dialogues on representation and recognition **199**

7.1 Dialogue 1: On the problem of representation 200

7.2 Dialogue 2: On theories of representation 205

7.3 Dialogue 3: On $\mathcal{S}$-isomorphism 207

7.4 Dialogue 4: On the Chorus of Prototypes 209

7.5 Dialogue 5: On the performance of Chorus 211

7.6 Dialogue 6: On representation and recognition in biological vision 214

7.7 Dialogue 7: On what has been left out of this book 218

A Measurement space **233**

A.1 The view space 233

A.2 The shape space 234

B Representation by distances to prototypes **239**

B.1 The Chorus Transform 239

B.1.1 Distance rank preservation by CT 239

B.1.2 Bounds on dimensionality reduction: the embedding theorems 240

B.1.3 CT as an implementation of Bourgain's method 242

B.1.4 Non-expansion of distances by CT 242

B.1.5 Preservation of Voronoi structure by CT 243

B.2 Representation by rank order of distances to prototypes 244

C Quasiconformal mappings and deformable shapes **247**

D RBF modules for object representation **249**

D.1 Radial Basis Function (RBF) approximation 249

D.2 Theoretical aspects of RBF module design 249

D.2.1 The infinitesimal displacement case 250

D.2.2 The finite displacement case 253

D.3 Practical aspects of RBF module training 256

D.3.1 Finding the optimal placement for each basis function 257

D.3.2 Tuning the basis-function width 259

E Vector quantization **261**

E.1 The generalized Lloyd (K-means) algorithm 261

E.2 Canonical Vector Quantization 262

F Multidimensional scaling **263**

G Performance of the Chorus system – additional tables **265**

H Trade-off between diagnosticity and invariance **269**