Title of Paper: **Towards a Transformational Approach to Perceptual Organization**

Names of Authors: (1) Douglas VICKERS, (2) Daniel J. NAVARRO, and (3) Michael D. LEE

Affiliations of Authors:
(1) Land Operations Division, DSTO
(2) Psychology Department, University of Adelaide
(3) Communications Division, DSTO

Contact Details of Authors:

Dr Douglas Vickers
Land Operations Division
Defence Science & Technology Organisation
PO Box 1500, Salisbury
South Australia 5108
Tel: +61 8 8259 7710
E-mail: Douglas.Vickers@dsto.defence.gov.au

Daniel J. Navarro
Psychology Department
University of Adelaide
Adelaide
South Australia 5005
Tel: +61 8 8303 5693
E-mail: daniel.navarro@psychology.adelaide.edu.au

Dr Michael D. Lee
Communications Division
Defence Science & Technology Organisation
PO Box 1500, Salisbury
South Australia 5108
Tel: +61 8 8259 5821
E-mail: Michael.D.Lee@dsto.defence.gov.au

Résumés:

Douglas Vickers gained an MA in philosophy from the University of Edinburgh and a BA, PhD, and ScD in psychology from Cambridge University. He is an Associate Professor of Psychology at the University of Adelaide, and is currently also a Visiting Research Scientist in the Australian Defence Science and Technology Organisation. His research interests are in visual perception and perceptual decision processes.

Daniel Navarro gained a BA in psychology from the University of Adelaide. He is currently engaged in postgraduate research at the University of Adelaide, under a joint scholarship from the Defence Science and Technology Organisation and the Commonwealth Government. His PhD thesis is concerned with categorization and the perception of similarity.

Michael D. Lee gained an MA in psychology and a PhD from the University of Adelaide. He is employed as a Senior Research Scientist in the Communications Division of the Defence Science and Technology Organisation and is also a Visiting Research Fellow at the University of Adelaide. His current research interests are in mathematical psychology generally and in analysing the complexity of theoretical models.

The presenting author is **Douglas Vickers**.

Towards a Transformational Approach to Perceptual Organization

Douglas Vickers[1], Daniel J. Navarro[2], and Michael D. Lee[3]
Defence Science & Technology Organisation[1,3], PO Box 1500, Salisbury,
South Australia, 5108, and University of Adelaide[2].
E-mail: Douglas.Vickers@dsto.defence.gov.au

ABSTRACT

A transformational approach to visual perception is presented, in which image structure is encoded in the parameters of those transformations that produce an output maximally symmetric with the current input. Results are presented for a computer program, dubbed SMART (Symmetry Maximizing Array using Random Transformations). SMART consists of a parallel array of independent symmetry detectors. Each detector attempts to find a transformation that maximizes the symmetry between the original and the transformed configuration. The weighted output of the detectors is collated in a connection matrix, which summarizes the image structure and provides a continuously varying measure of relative symmetry. The program is applied to constrained and random arrays, Glass figures, and the detection of hidden symmetric targets. More general implications are briefly discussed.

Most recent attempts to develop a general explanatory framework for phenomena belonging to the intermediate levels of visual perception can be classified into one of two broad categories. *Likelihood* approaches argue that image elements are organized by unconscious inference processes into the *most likely hypothesis* concerning their source (e.g., Albert & Hoffman, 1995; Knill & Richards, 1996). *Simplicity* approaches argue that perception organizes image elements so as to provide the *simplest description* or encoding of the source of that stimulation (e.g., Hatfield & Epstein, 1985; Van der Helm & Leeuwenberg, 1996).

Although both approaches provide plausible accounts of many phenomena, each is vulnerable to criticism. A problem for the likelihood approach is that assumptions required to assign likelihoods to hypothetical situations giving rise to an image tend to be elaborate and tenuous. On the other hand, the simplicity approach requires debatable assumptions concerning the basic elements of a description and the process by which descriptions are constructed. More importantly, neither approach suggests a specific mechanism by which their optimizing principles might be implemented.

SMART (Symmetry Maximizing Array using Random Transformations)

We propose an alternative, *transformational* approach, motivated by the obvious importance of symmetries and transformations in the recovery of object information in visual images (Barnsley, 1993; Mundy & Zisserman, 1992; Tyler, 1996). This approach follows previous theorizing (Palmer, 1999), but differs in that it attempts to explore specific mechanisms for maximizing symmetry.

As a first step, we have developed a program in Matlab, dubbed SMART (Symmetry Maximizing Array using Random Transformations). Like someone trying to solve an anagram, the program begins by subjecting image elements to multiple random transformations. Symmetries produced by this process then select those transformations that best capture any structure in the image.

Figure 1 illustrates the structure of the program. This consists of S symmetry detectors, operating independently of one another. Each detector inputs the coordinates of the set of points, P, normalized to be between 0 and 1, and subjects this set to some transformation, to yield a transformed set, P_t. Allowable transformations are currently restricted to combinations of rotations and vertical and horizontal translations, randomly selected from normal distributions, centered on zero. Standard deviations, σ_r, σ_v and σ_h, are typically set to values of $\pi/16$, 0.05, and 0.05, respectively.

For each point, i, in P_t, the program finds the point in P that is closest to it, and records the distance, d_i, between these points. (To prevent null transformations, the restriction is imposed that point i in P cannot be considered for point i in P_t.) The distance, D, between the two point sets is evaluated as the sum of the individual inter-point distances, $D = \Sigma d_i$. The program then uses a hill-climbing algorithm to find a transformation, t, and hence the point set P^*, that corresponds to a local minimum of D. The coordinates of the two point sets, P^* and P, are then compared. For each point i in P^*, the distance d_i to its nearest neighbor in P is calculated. If this distance is less than a predefined tolerance, t, then a mapping $\{i, j\}$ of that pair of points is recorded in a connection matrix, C.

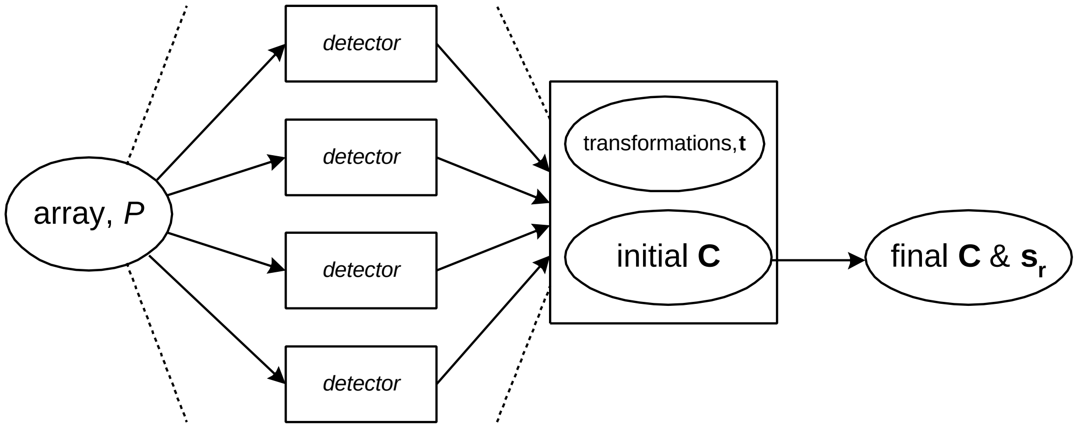


Figure 1. Information flow diagram of the processes involved in the SMART program.

The matrix C starts as an NxN matrix of zeroes, where N is the number of points in P. The output of each detector then modifies C as follows. Each mapping $\{i, j\}$, made by the detector, corresponds to two entries in C (i.e., to c_{ij} and to c_{ji}). Both of these entries are increased by x, where x is the total number of mappings made by that detector. This weighting mechanism allows the best transformations to dominate C.

After all detectors have modified C, the matrix is then normalized, as follows. The maximum value an entry, c_{ij}, can have occurs when all S detectors make the maximum of N mappings and the particular mapping $\{i, j\}$ is made by every detector. Since this value will be equal to NxS, we normalize C by dividing each entry by NxS. The resulting normalized matrix constitutes a cumulative record of the major point symmetries discovered by the detectors. This record can be used, together with an Sx3 matrix of the associated transformations, to identify those transformations that maximize invariance in the image. The program can then be set to display only those mappings that contribute more than a certain threshold value, l, to the matrix.

The SMART program also provides a measure, s_r, of the relative symmetry of the array, according to the following rationale. If no detectors map any points, then we want s_r to be 0. Conversely, if all S detectors map all N points, then we want s_r to be 1. A detector that maps N points adds N to each cell in C that corresponds to a mapping, and hence adds a total of $2N^2$ (since each mapping corresponds to two entries). Thus, the upper bound on $\Sigma_{ij}C_{ij}$ is $2N^2$xS. After normalization, this total becomes $2N^2$xS/(NxS), or $2N$. We define s_r as $\Sigma_{ij}C_{ij}/2$N. This is 0 when no mappings are made and 1 when all detectors find a perfect symmetry.

Some examples of SMART analyses

Despite its simplicity, SMART is effective at identifying structure in point arrays. The following examples illustrate the range of situations in which the program is successful in detecting structure and in simulating human performance.

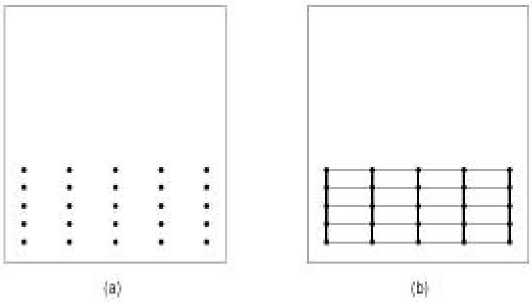
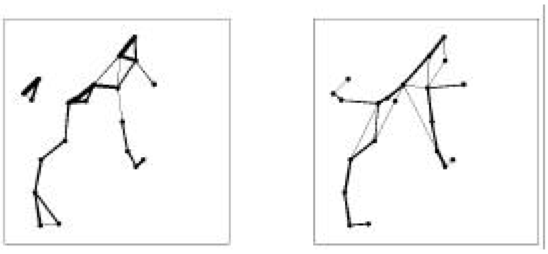


Figure 2. Illustration of how SMART detects simple translational symmetries and captures the Gestalt principle of organization by proximity as a result of the normally distributed values for initial transformations. For this analysis, the number of detectors, S = 20, the drawing threshold, l = 0.05, the mapping tolerance, t = 0.025, and the standard deviations for rotations, σ_r, and for vertical and horizontal translations, σ_v and σ_h, were $\pi/16$, 0.05, and 0.05, respectively.

Gestalt structure. As illustrated in Figure 2, if the program is applied to a regular array of points, which are naturally perceived as organized into columns rather than rows, it outputs a diagram of this structure. Thicker vertical lines indicate that the program has accorded a greater weight to vertical than to horizontal translations, in agreement with human perception. However, the preference for structure based on proximity is not explicitly built in, but follows from the assumption of normally distributed parameter values for the transformations carried out by the detectors.



(a) (b)

Figure 3. Structure detected by SMART in the constellation *Perseus* (a), with a summary of its representation in seven star atlases (b). (Thicker lines indicate more frequently connected points.) For this analysis, S = 20, t = 0.1, l = 0.05, and σ_r, σ_v and σ_h were π/8, 0.1, and 0.1, respectively.

As illustrated in Figure 3, SMART is also capable of detecting structure in less regular arrays, such as the well-known constellation, *Perseus*. So far, we have not directly compared this capability with human performance. However, the summary data from seven star atlases, shown in Figure 3b, suggests that some configurations at least may be captured quite successfully.

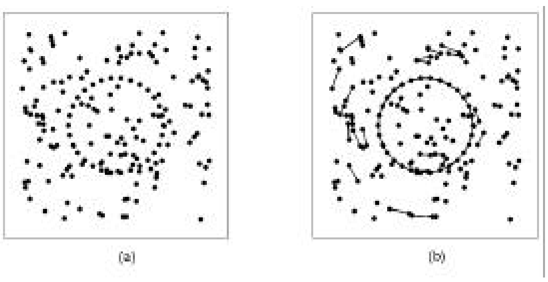


Figure 4. Illustration of how SMART can detect a regular configuration, or target, hidden in noise. The target consists of a circle, represented by 20 equally-spaced points, embedded in 150 randomly distributed points. For this analysis, S = 20, t = 0.025, l = 0.2, and σ_v and σ_h, were π/8, 0.1, and 0.1, respectively.

Hidden figures. More generally, the program is capable of detecting structure in any regular, semi-regular or irregular array. In addition, as illustrated in Figure 4, it can detect complete or partial regular structures embedded in random arrays. It is worth emphasizing that SMART has not been programmed to detect any particular structure in this array. This capability, in particular, may have some industrial or medical applications.

Glass patterns. Figure 5 shows two examples of so-called Glass patterns (Glass, 1969). These textures are generated by taking a uniform random spatial distribution of dots and superimposing a geometrically transformed copy of the set. Such textures are perceived as having a clear structure, consisting of dot pairs (or dipoles), locally aligned in the direction of the transformation used to generate them. Glass patterns are of interest because, in order to perceive structure in such textures, it seems necessary to suppose the visual system must be solving some form of correspondence problem. This is not straightforward because perceptual grouping based on proximity, for example, may operate against an organization based on (say) rotational symmetry.

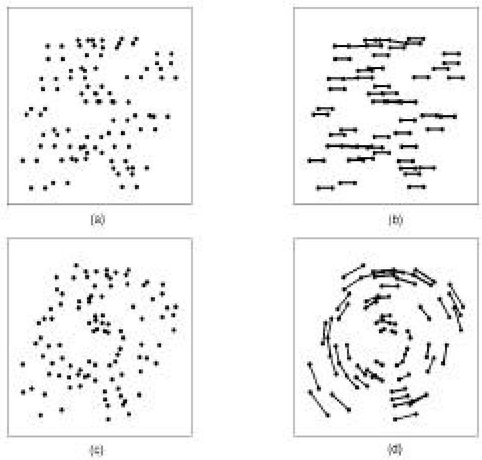


Figure 5. Glass patterns generated by (a) horizontal translation and (c) rotation and the corresponding mappings (b) and (d) detected by the SMART program. Parameter values were as for Figure 2.

As can be seen from Figure 5, Glass patterns generated by translations or rotations pose no problem for the SMART program. Although the program does not currently incorporate dilations among the set of permissible transformations, it is not difficult to extend the program in this way. An advantage of this approach is that it detects global structure without the need for higher-level processing on the orientation of point pairs.

Relative symmetry. According to the traditional conception of symmetry, a figure either has a particular symmetry or it does not. However, most naturally occurring structures are characterized by imperfect symmetries that seem to vary in a continuous way. The SMART program provides a continuous measure of *relative symmetry*.

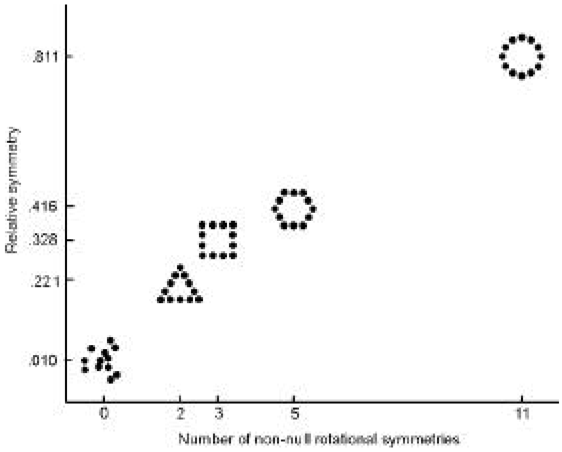


Figure 6. Depiction of changes in the relative symmetry measure, s_r, with increasing rotational symmetry. Parameter values for this analysis were S = 100, t = 0.025, and σ_r, σ_v and σ_h were $\pi/16$, 0.05, and 0.05, respectively.

As illustrated in Figure 6, this measure is a direct, monotonic function of the number of rotational symmetries in regular *n*-gons. More generally, however, the measure can be applied to arrays of any degree of complexity or irregularity, including many that would be characterized by human observers as 'more or less' symmetrical. This means that SMART has the potential not only to capture this aspect of perception, but also to generalize the application of approaches based on symmetry, and make possible their systematic investigation.

Transformational approach in general

Early explorations with this transformational approach suggest that it has the potential to provide a simplified, but workable, model of certain aspects of human visual perception (Vickers & Preiss, in press). Because it is a generative theory, and implies some understanding of the development or construction of objects, this approach shifts the traditional boundary between perception and cognition. For the same reason, the transformational approach has the potential to account for our perception of the 'process history' of an object (Leyton, 1992) and for the 'representational momentum' that allows us to extrapolate to its future state (Freyd, 1987). For the same reason also, the transformational approach suggests that perceptual information may be remembered in a more dynamic form than has been generally supposed - one that is directly related to the way in which such information can be regenerated (Vickers & Lee, 1997).

REFERENCES

Albert, M.K., & Hoffman, D.D. (1995). Genericity in spatial vision. In: R.D. Luce, M. D'Zmura, D. Hoffman, G.J. Iverson, & A.K. Romney (Eds.), *Geometric representations of perceptual phenomena* (pp. 95-112). Mahwah, NJ: Lawrence Erlbaum.

Barnsley, M.F. (1993). *Fractals everywhere*. 2d ed. San Diego, CA: Academic Press.

Freyd, J.J. (1987). Dynamic mental representations. *Psychological Review, 94*, 427-438.

Glass, L. (1969). Moiré effects from random dots. *Nature, 243*, 578-580.

Hatfield, G.C., & Epstein, W. (1985). The status of the minimum principle in the theoretical analysis of visual perception. *Psychological Bulletin, 97*, 155-186.

Knill, D.C., & Richards, W. (1996). *Perception as Bayesian inference*. New York: Cambridge University Press.

Leyton, M. (1992). *Symmetry, causality, mind*. Cambridge, MA: MIT Press.

Marr, D. (1982). *Vision*. San Francisco, CA: Freeman.

Mundy, J.L., & Zisserman, A. (1992). *Geometric invariance in computer vision*. Cambridge, MA: MIT Press.

Palmer, S.E. (1999). *Vision science: Photons to phenomenology*. Cambridge, MA: MIT Press.

Tyler, C.W. (1996). *Human symmetry perception and its computational analysis*. Utrecht, the Netherlands: VSP.

Van der Helm, P.A., & Leeuwenberg, E.L.J. (1996). Goodness of visual regularities: A non-transformational approach. *Psychological Review, 103*, 429-456.

Vickers, D., & Lee, M. (1997). Towards a dynamic connectionist model of memory. *Behavioral and Brain Sciences, 20*, 40-41.

Vickers, D., & Preiss, A.K. (in press). Transformational analyses of visual perception. *Proceedings of the Twenty Second Annual Meeting of the Cognitive Science Society.*