On the Visual Perception of Shape

Analysis and Genesis through Information Models

Frederic Fol Leymarie

Goldsmiths College, Computing, University of London, U.K.

"Vision is continuous creation of meaning from images." — R. L. Gregory [15]

Abstract. Shape is the structure of a localised field constructed in relation to an object, which *e.g.* consists of geometric entities in association with each sample of the object being scrutinized. To identify or characterise shape, we need to "probe" the field [27,35]. In this paper I consider the visual perception of shape via "information models" defined as simultaneously usable for both tasks of: *(i) Genesis, i.e.* the creation of an internalised world construct, part of a "presentation" [28], and *(ii) Analysis*, *i.e.* stimuli processing to produce perceptual cues to steer the generative engine as well as for verification of the generated percepts.
Keywords: Shape, Information models, Presentations, Medial representations, Shape grammars, 2D patterns.

1 Introduction

Perception involves constant choice, ceaselessly deciding which meaning to impose on the elements of sensory data.[1] Why might this "strategy" have evolved? For any given set of sensory data, there are generally multiple interpretations possible, but the organism must be able to rapidly come to a perceptual conclusion if it is to avoid being attacked, trapped, or put into a dangerous or uncertain situation. There may not even be a single optimal interpretation of the stimulus, as seen in ambiguous patterns like the celebrated Necker cube [13]. And yet the organism needs a perceptual system capable of making decisions constantly, even if those decisions are simplistic approximations, and sometimes outright wrong; thus the perceptual system must be continuously revising and re-evaluating its interpretations, as illustrated in the bi-stable Necker cube. Necker's line drawing need not be a 3D object itself, but must therefore be a construct of our mind, correlated with sensory evidence.

Although we continuously sample the visual environment, we do so in general with small visual saccades, fixating mostly to update visual knowledge, to analyse slight changes, possibly to compare features (of a painting, a face). But, during saccadic

[1] I refer to "data" in its usual meaning in communication theory: the transfer of information between an emitter and a receiver; *e.g.* the emitter may be (sources of) light bouncing off objects in a scene or signals generated by the brain when dreaming, while the receiver is the visual area (or "interface") of the brain which encompasses (at least) the eyes, the optic nerves, the LGN (lateral geniculate nucleus, the "relay" centre from the retinas to the visual cortex), V1 (the primary visual cortex) and others areas of the nervous system directly involved in "visual processing". But, notice that data is only involved in part of what perception is about: data stimulates the mind in its creative quest and helps it quickly produce an answer.

eye movements, the entire visual field (actually the entire perceptual field — including sounds, haptics, heat sensing) remains stable [51, pp. 464-5]; we have a clear view: there is no gap, hole, stitch, vibration, or noise in our (conscious) understanding of the scene. This I argue (with others) is possible if perception is a *creative* process: rather than just being built up from the sensory data, the perceptual experience involves an internal creation which is "colored" [21,16,28], constantly updated or verified using our sensors (and their incoming data). Such an internal creation by the mind of a view of the world is called a *presentation* [28].

The theory of shape perception I explore is to be based on such a theory of the mind. "Shape" is not strictly and directly only a set of physical properties derived from objects in the world but is foremost a proposal of the mind, a presentation, which is correlated with the physical cues available to our senses.[2] Hence shape cannot simply be derived from affordances of object out in the world, but can rather only be understood via the interface (here visual) that our mind controls [19].

2 Shape percepts: Analysis & Genesis via Information Models

Main Hypothesis: The mind requires as one of its key building blocks sets of *information models* simultaneously usable for both tasks of: *(i) Genesis, i.e.* the creation of an internalised presentation (of meaningful percepts), and *(ii) Analysis*, *i.e.* stimuli processing to produce perceptual cues to steer the generative engine as well as to verify the generated percepts. That the *same* information model is needed for both genesis and analysis is a consequence of the constant "conversation" that takes place between the two tasks: it leads to efficiency for a biological system that needs rapid decision making, right or wrong.

What is an Information Model (IM)? In essence it is what embodies *meaning*. Loosely speaking, it is a *description* of an object, resulting in a visual percept (of a shape or set of shapes in the discussion here) in terms of qualitative properties, quantitative features (*e.g.* geometrical), relationships (*e.g.* topological), associated functions and contexts.[3] In order to illustrate and put in practice this concept, I will start using a natural language description of shape percepts and then propose a possible computational implementation. I will restrict myself in this contribution mainly to 2D percepts as "shapes," but what is exposed extends to 3D objects and their associated shape.

Definition 1. ***Shape*** *is the structure (the organisation, arrangement of parts) of a localised field constructed "around" an object. This field consists* e.g. *of geometric entities — such as curvatures, singularities (of some appropriate mappings), gauge figures (*e.g. *Dupin's indicatrix) — which may exist in association with each sample of the object being scrutinized, or are observed for regions, parts or the entirety of the object's*

[2] This applies also to other perceptual categories, such as color percepts.

[3] I am borrowing and adapting the concept of IM from conceptual modelling and software engineering, where it is defined at its most general level as a set "of concepts, relationships, constraints, rules, and operations to specify data semantics for a chosen domain of discourse; it can provide sharable, stable, and organized structure of information requirements for the domain context" [31].

trace. In order to identify or characterise shape, we need to (operationally) "probe" the field (adapted from [27,35]).

I therefore cast this definition of shape in the framework of Information Models and explore probes, or operators, which are useful for both genesis and analysis tasks.

2.1 An example of an Information Model for shape: Descriptions of 2D patterns

As an example of the visual perception of shape, let us consider simple 2D percepts like a triangle, a corner, a rectangle or a square.

Corner: the region of an object made of boundary segments meeting at a sharp end which is typically located at the intersection point of the pair of segments.
Rounded corner: a corner with a rounded (not so) sharp end.
Triangle: a 2D object, typically with an explicit closed outline, made of three (usually sharp) corners. These corners are often linked by straight line segments.[4]
Quadrilateral: a 2D object, typically with an explicit closed outline, made of four (usually sharp) corners. These corners are often linked by straight line segments.
Rectangle: a quadrilateral with right angles at each of its four corners.
Square: a rectangle with equal length boundary segments or with equally distanced corners along the boundary.

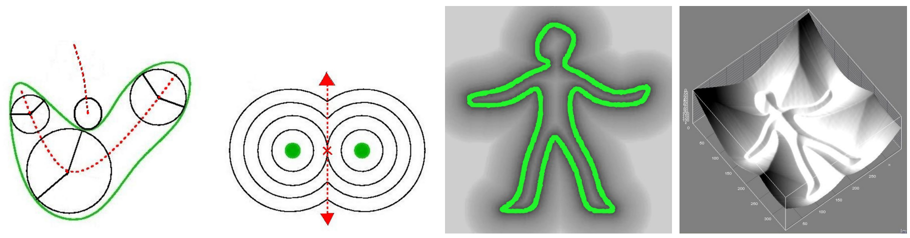
Each of the above shape description combines boundary features with some interior or regional properties: *e.g.*, a corner is the sharp(er) interior end region delimited by a meeting pair of lines tracing contour segments.[5] A triangle is a closed envelope (a solid shape in technical terms: *i.e.* we can fill it with a fluid which cannot (or with difficulty) escape that interior region) delimited by three corners. A shape may be described in terms of another: a square is a more symmetric rectangle; a rectangle is a straighten parallelepiped, *etc*. When specifying a deformation to relate a more complex shape to a simpler or more symmetric one, we can follow *e.g.* a dynamic principle of least perceptual effort (or tension reduction when referring to Gestalt psychologists [2]). A high level study of dynamic form (in terms of group theory) is to be found in Michael Leyton's generative shape theory [38].[6]

2.2 A potent computational substrate for an Information Model of shape: Medial representations of 2D patterns

I explore now a potential computational substrate based on Harry Blum's medial axis framework [8] (Fig.1), which can support or even replace[7] natural language descriptions like the ones above for simple shapes. Part of my interest in this computational

[4] I allow non-straight (or non-Euclidean) boundaries so that bending (or geodesic/Riemannian) triangles are accounted for. I also allow for open topologies, to account for potential illusory contours, such as provided by the Kanisza examples (Fig. 2).
[5] A different (operational) description is likely to lead to a different implementation of the IM.
[6] Leyton's work gives a view on shape via operators equipped with control structures and feedback mechanisms.
[7] An IM for shape can support natural language descriptions, but this need not be a necessary requirement, especially if a theory of shape perception is to be generalised to other species than humans.

framework in the present context is that it relates contour features with object regions and is both used for analysis [25] and genesis [6,5] of forms. It also appears powerful enough to study at least some of the known (shape-related) perceptual illusions such as the Kanizsa-like ones (Fig.2) [25]. The medial axis can be defined in a number of ways, such as (i) the loci where wavefronts initiated at boundaries collapse or meet in a simulated field (or "grassfire") propagation, (ii) the ridges of a distance field map — where the distance is computed away from the object's boundary and plotted as a height map, or (iii) the loci of center of maximal balls osculating two or more contour points, [8] (Fig.1).



(a) (b) (c) (d)

Fig. 1: Adapted from [49, Fig.3]: The Medial Axis of Blum: (a) as the loci of centres of maximal contact disks; (b) for two sample points as the locus of meeting Euclidean wavefronts, where the (central) cross indicates the initial shock when the fronts first meet, and the arrows indicate the direction of growth or flow of the medial axis segments; (c) derived from a distance map computed for a humanoid outline, as the ridge lines of the corresponding height field, (d).

While the original ideas of Blum emerged in the 1960's, a more recent development for this computational framework takes the form of shock graphs (in 2D) and medial scaffolds (in 3D) [25,36,44] for which it is shown that the medial axis is somewhat redundant as a diagrammatic model of distance-based symmetries, such that the critical points of the radius (or distance) "flow" along its trace are sufficient to obtain a compact and complete view of Blum's symmetries.[8] I will not discuss here all the properties of

[8] Other definitions have been proposed since Blum's early ideas, including most notably: (i) The *Process Inferring Symmetry Axis* (PISA) of Michael Leyton [38]: an extension of the medial axis putting the emphasis on growth and a process-based view of shape — shape is understood as the history of the deformations likely to explain it; (ii) the *M-rep* of Stephen Pizer *et al.* [44]: the medial axis is approximated via primitive sub-shapes (functioning as parts) which can explain the segmented data — typically the samples of outlines obtained from an image processing step — and for which the medial representation is explicitly known; such M-reps are meant to explicitly re-create an approximation of the original object as the concatenation of the fitted primitives; (iii) the *annulus-based medial fields* of Kelly and Levine [24], Kovacs *et al.* [29], van Tonder *et al.* [48]: symmetries are captured for a disk with a thick boundary — any outline samples falling in the annulus region count towards a symmetry — making the

such medial models (refer for example to Blum's main thesis [8], my previous study on shape [35] or to the monograph edited by Siddiqi and Pizer [44]); again, one important reason why such medial representations are suitable to build IMs is that they can be simultaneously used for both purposes of analysis and genesis. Let us then consider a possible mapping between a shock graph and a natural language descriptions for our above examples of corner, triangle and quadrilaterals.

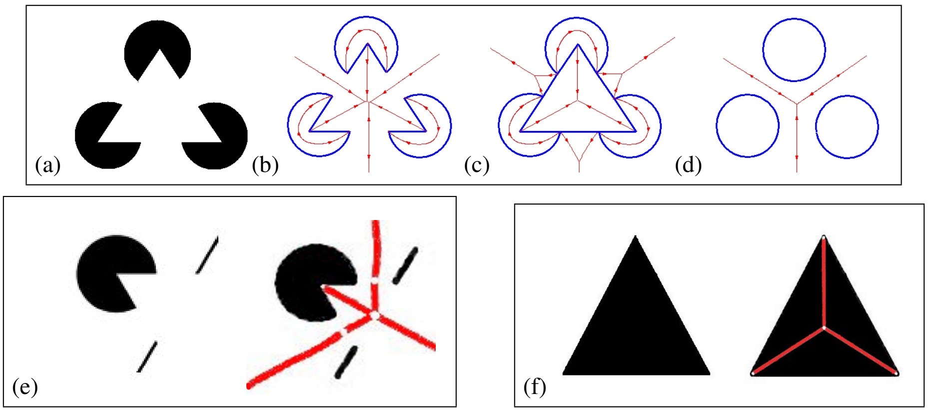


Fig. 2: **Top line:** adapted from [25]: (a) the Kanizsa triangle [22]: note that the lightness of the "triangle" perceptually exceeds that of the paper or screen forming a background, which illustrates one type of our mind's constructive processing: this additional brightness is only available in the observer's presentation, not in the physical world. (b) The shock graph arising from the edge map directly recovered from the Kanizsa stimuli, where arrows indicate flow; this edge map can be transformed (edge elements removed, gaps bridged, and so on) such that any transformation sequence of the edge map is implemented by a corresponding sequence of symmetry transforms on the shock graph. The sequence used here recovers the illusory triangle at one depth, (c), and completes the "pacmen" as disks at another, (d). NB: The selection of the proper transform sequences — *i.e.* how to navigate in the shape space of deformations (here: symmetry transformations) — remains a difficult task in this bottom-up framework. Other sequences, with similar deformation costs, lead to very different unintuitive results. **Bottom line:** Examples of shock graph computations (e) for a 2D corner extracted from a Kanizsa-like illusory percept, and (f) for a (usual, equilateral) triangle as a reference. Notice that the white dots along the shock graph indicate the extrema of the flow along its trace (min, max and saddles). Notice also that a similar additional brightness effect as observed in (a) is available to the observer's presentation on the left in (e), however, with no well defined boundary away from the corner (or pacman), illustrating that illusory contour presentations do not require solid shape percepts, merely open ones.

representation non-unique, but more robust to perturbations in the data; (iv) the *full symmetry set*: captures a richer set of symmetries for a given outline by letting wavefronts propagate and cross each other multiple times [12,44].

Corner: a corner is indicated by a shock graph branch which emerges from a high curvature point on the boundary; that branch puts in symmetry relation two bounding contour segments; the local region for which the distance field is computed in relation to the shock graph captures the size, width, elongation and propensity to "compress out fluid" (or distance flow field) away from the corner region to other (shape) regions; the higher the curvature the closer the shock branch end gets to the contour and touches it for a break of curvature at a sharp end. Blum's medial axis can be naturally applied to partial boundary segments, or contour with holes, which makes it an interesting candidate to study Kanisza triangle like illusions and other patterns with occlusions, noise, gaps (Figures 2 and 3).

Triangle: a triangle is indicated by a shock graph with three branches each emerging from one of the three triangle corners; each branch relates to the two sides of each corner; the shared relationship between each boundary segment by pair and as a linked triplet is explicit; the triangle border (and interior region) is always available from the graph by sweeping primitive functions along each branch (in the form of disks of smoothly varying prescribed/encoded radii). Note that for retrieval purposes from a shape-encoded memory, the mapping of the shock graph structure for a triangle is readily available in the Kanisza triangle stimuli.[9] Also, different radius functions permit to capture bends and bulges (*e.g.* for the deformed Kanisza stimuli of Fig.3, b&d).

Quadrilaterals: these can be studied similarly to the triangle case. The geometry of the shock graph distinguishes most of the variants (general quadrilateral, parallelepiped, rectangle) while special topological events, such as the higher symmetry of a square or of a diamond, are made explicit in the (simplified) branching structure. A similar discussion on illusory percepts like the Kanisza triangle applies to quadrilateral configurations.[10]

IMs with implementations carried out by shock graphs for 2D patterns (or medial scaffolds for 3D ones [36,9]), can potentially capture large classes of 2D (and 3D) visual shape percepts. We may restrict our shape space to the study of objects that are reasonably regular: not necessarily with smooth boundaries everywhere — the common form of mathematical regularity — such as outlines with discontinuities (in curvature or with gaps), but also not arbitrarily complex, as can be generated by fractal and some stochastic processes.[11] This restriction allows us to address the issue of simplicity. For regular objects we expect medial representations such as shock graphs and scaffolds to be relatively simple, consisting only of sparse directed graph structures constructed from nodes at singularities of the medial flow and their connectivity. The genesis or the analysis can be performed at varying scales, and for various subparts. The genesis can be performed in various ways, and in particular by inverting the process used for the

[9] This is typically performed in computing via graph matching techniques developed to deal with partial cues, *e.g.* using edit matching [43].

[10] A detailed study of 2D illusory patterns with medial representations is beyond the scope of this contribution.

[11] Medial representations can be used to study and generate fractals and some stochastic processes; thus our restriction may not be critical or even necessary if we want to generalise the proposed approach.

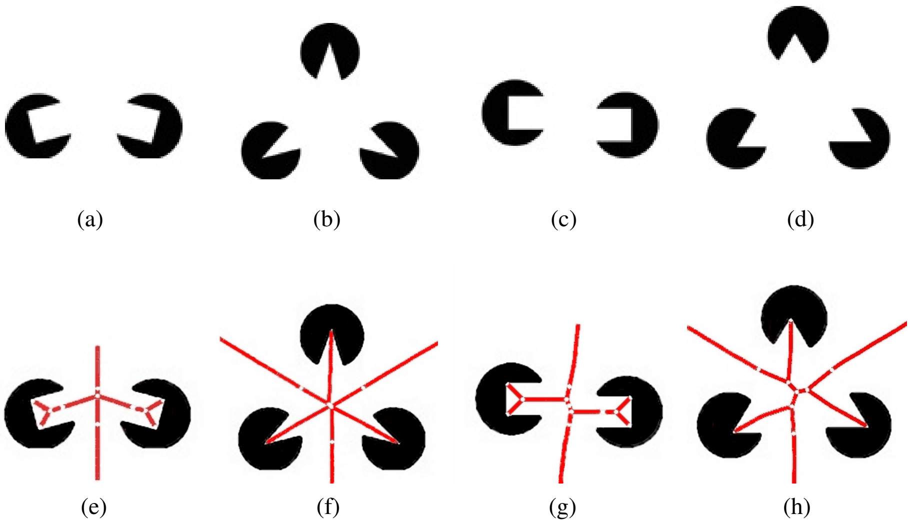


Fig. 3: **Top line**: after Lehar [32, figure 1]: (a) and (c): the stimuli of Kellman & Shipley [23] used to study illusory contours as a function of (a) a bending misalignment, and (c) a shearing misalignment; (b) and (d): corresponding equivalent Kanizsa figures. **Bottom line**: corresponding shock graphs; in (e,f) a bending misalignment is made explicit from shock graph geometry and associated radius function flows, while in (g,h) a shearing misalignment is noticeable from the shock graph in its topological splits (new branch segments introduced in the central area).

analysis. For example, from a shock graph trace, a sweeping function can be used by varying the radii of discs, or a reverse distance transform can be computed.

Higher order symmetries, in particular those in correspondence with medial flow singularities, organise the shape space in regions providing "attractors" around which deformed versions can be explicitly related [25].[12] For example, the square organises rectangles and other quadrilaterals, under various projections, "around it;" and similarly does the circle for ellipses, ovals and other slightly deformed disks. This extends to 3D with cubes or spheres. This analysis naturally encompasses the tubular shapes of Gaspard Monge (classical geometry), the geometric primitives (geons) of Biederman *et al.* [3], and more flexible "generalised cylinders" (sweep functionals) favored by Binford, Marr and others [41].

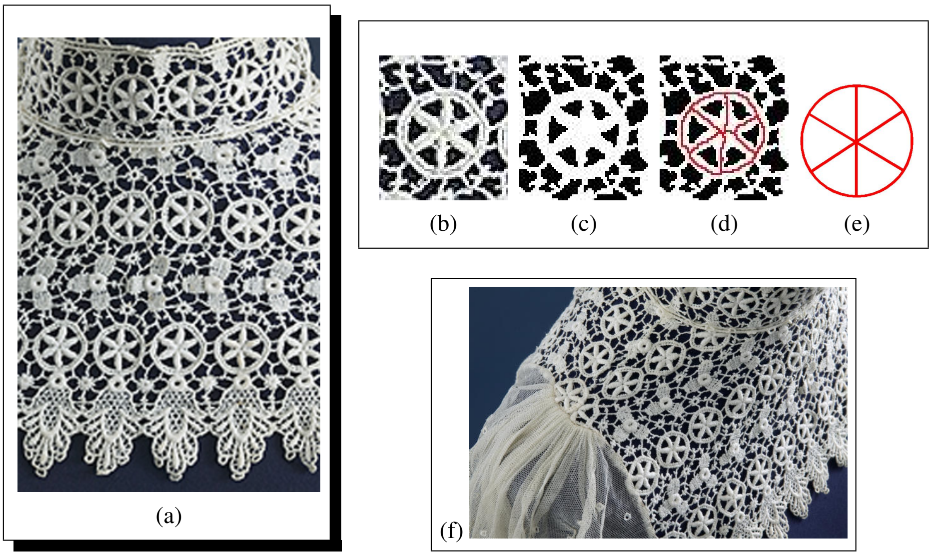


Fig. 4: Example of a pattern analysis via an attractor-based shock-graph recovery. (a) Part of a collar (lace fabric; from the Constance Howard textile collection at Goldsmiths, University of London). (b) Selected/extracted patch containing a pattern of interest. (c) Figure-ground segmentation by thresholding of the image patch. (d) Recovery of the main shock graph structure (analysis). (e) Corresponding highly symmetric graph target: an attractor in the shape space. This attractor has the potential to capture either a natural language description or symbolic enunciation of the geometrical and topological nature of the class of objects weaved into the textile. (f) It can then be used to find other images in this or other collections where the same or similar pattern(s) exist.

[12] This goes against the traditional view favored in applied mathematics and physics: high-order symmetries are ruled out as non-significant (*i.e.* not occurring in nature) since a slight perturbation of the data or the stimuli breaks these away. The psychological realm of perception needs not follow the same constraints if it can fully control its generated shapes.

Families of patterns can thus be captured by their representative *attractor* — defining a region of the shape space which can be associated to a class of shapes; *e.g.* all the circumscribed star-shape patterns found in the piece of lace in Fig.4 are captured in one class by the attractor in Fig.4.(e).

Beyond strictly the domain of "simple" pattern description, we can seek IMs able to deal with more complex visual contexts. One such important computational substrate can be found in early and more recent work on shape and image grammars.

2.3 Information Models & Shape Grammars in Biology, Architecture & Vision

Generative grammars expressible in a computational framework have a significant recent history (post-WWII), starting with Chomsky's theory of language [10] which introduces the notion of transformational generative grammar based on powerful rewrite rules providing explicit and formal descriptions. This approach later transited from linguistic to tackling shape problems. To my knowledge, this was first accomplished by Lindenmayer *et al.* [39,42], in particular to tackle in biology the modeling of plant structure as well as their growth patterns. Later other shape grammars were proposed in architecture (Stiny *et al.* [45]) and in vision and related applications (Leyton [37,38]). Perhaps the most successful recent practical application of grammars is to be found in Computer Aided Design [11,20]. In general, such (shape) grammars are presented either in a generative context or in an analytical one, but they can be designed to simultaneously serve both purposes, making these interesting candidates as IMs that can directly link a semantic description of a perceptual object with a computational implementation.

The first significant use of generative processes in *computational architecture research*, referred to as "process grammars", was introduced in the work of George Stiny *et al*. [45].[13] Such grammars consist of a database of rules and a generative engine and represent a form of expert system where the generative engine checks the existing created geometry for conditions that match the left-hand side of the rules. Defining an appropriate set of rules is usually very challenging, is building type and architectural style dependent, and that set of rules can rapidly grow. In more recent work, the resulting shapes are constrained by introducing parametric controls to the right-hand side of the fired rules, as illustrated in the work of Mueller, Wonka *et al.* [40]. Theirs is an extension of Stiny's shape grammars and includes split rules, some "merging" capacity (creating joints/snapping constraints, evaluating intersections), some control mechanisms (of the grammar), scaling, tiling and basic Euclidean transformations, context sensitive shape rules, and stochastic (placement) rules. It also relies on existing computational geometry algorithms, such as straight skeletons to derive plausible roofs,[14] and specialised (but well-known) geography-based and computer vision (photogrammetry) mechanisms to extract relevant outlines and (photo)textures from imagery: (i) for footplates and street layouts from aerial imagery, (ii) from facades to retrieve regular layouts and textures (of windows, doors, decorations, gutters) [46].

[13] The tradition of defining and using rule-based systems in architectural design is ancient; two of its most well-known champions being Vitrivius (Rome, 1st century BC) and Palladio (Renaissance, 15th century) who produced treatises on the subject.

[14] Straight skeletons are yet another form of medial representation.

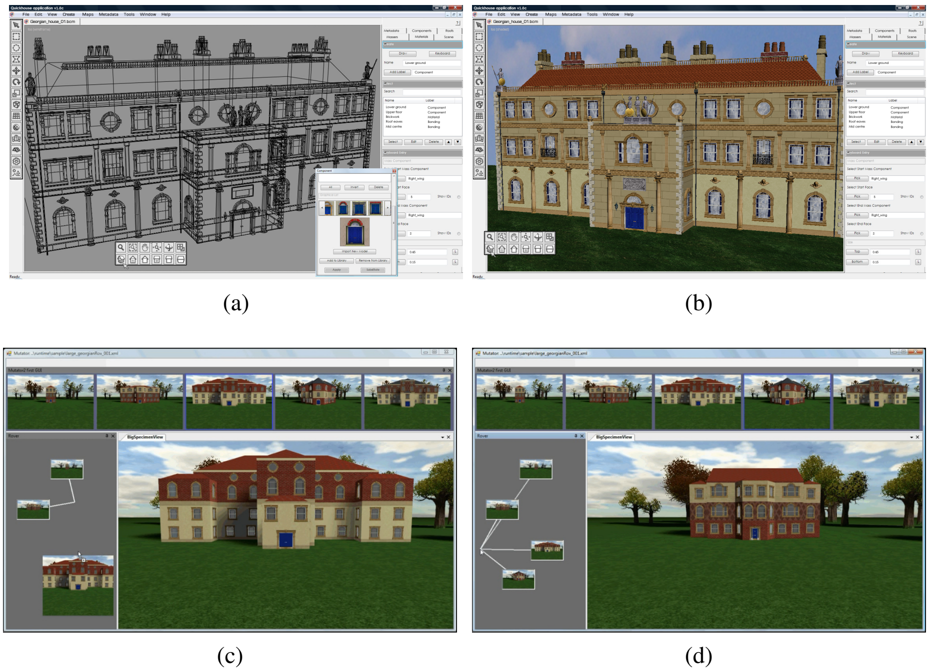


Fig. 5: Example of the use of a Building Information Model (BIM) part of a system being developed by the Mutators Research Group at the University of London. In (a) and (b) are shown two views of a Georgian style building designed using an ontology including definitions of footplates, walls, floor levels, rooftops, windows, doors, decorative objects, material properties, *etc.* In (c) and (d) are shown two examples of generated buildings obtained by a "genetic mixing" derived from the ontology-based descriptions of multiple existing styled buildings seen as their "parents": in (c) for a pair of parents and in (d) for a quadruplet of parent buildings. This work is a continuation of previous projects initiated by Latham and Todd in the early 1990's [47] and pursued at the University of London in the last few years [30].

In the last decade, the *architectural* and *engineering professions* have started to move from an electronic drafting process towards a parametric modelling of objects. Building Information Models (or BIM) as they are referred to in the architectural and built environment fields permit the parametric modeling of all objects (walls, doors, windows, floorplans, *etc.*) and their interactions, combining shape generation of buildings and their component, together with a knowledge base in the form of an architectural ontology [1,17]. This is made possible by defining functional structural relationships between shape elements and managing these in a database or memory structure. Most implementations of the BIM philosophy follow the IFC (Industry Foundation Classes) which insures greater interoperability between disciplines and software packages. The IFC model represents architectural shapes (*e.g.*, doors, walls, tiles, etc.) as well as more abstract concepts such as spaces (*e.g.*, floorplans), construction costs, activities/functions, schedules/dynamics and so on. All entities can have associated properties, such as materials, name, geometry, textures, relationships, lifetime, and so on (Fig. 5).

In recent years Zhu and Mumford [52] have revisited the use of grammars in *vision* to address the problem of dealing with large amounts of possible object categories — for example, according to work by Biederman, at least 3,000 shape categories can be derived from the English dictionary, where each category can cover large amounts of variants [3,4]. Such grammars represent the hierarchical nature of scenes into objects, parts, primitives, their relationships and relative scales. They also address issues of spatial and functional contexts. Zhu and Mumford's grammar is implemented as And-Or graphs where an Or-node corresponds to alternative sub-structures, while an And-node corresponds to parts. This model is used both for the analysis of images and genesis of novel configurations (for example of an object like a chair, a clock or a car and its parts).

3 Discussion

Grammars in vision have four important objectives [52]: (a) to offer a common framework for visual knowledge representation and object categorisation, (b) to offer scalable and recursive top-down/bottom-up computations, (c) to function from small sample learning and be able to generate novel configurations, (d) to offer a mapping between the visual vocabulary at multiple levels and symbols and stimuli, therefore solving the semantic gap. With respect to natural language, there are three main additional issues: (i) the loss of a 1D (time) ordering, (ii) the presence of image scaling and (iii) the presence of a wider range of irregular patterns (than found in speech signals).

Medial shape operators and shock graphs in particular are one of the main streams of research which can support image grammars [52]. When in the form of shock graphs (2D) or medial scaffolds (3D), a medial shape operator can *explicitly* be represented via a *neural circuitry* (brain machinery) [25]. It can also be represented *implicitly* via global brain (wave) functions, as proposed by Blum [7], and recently revisited by Lehar [33,34]. Global brain wave functions exploit the *field-like* structure which supports medial axis computations, *i.e.*, the directed distance field — resulting from either an *analysis* like in the grassfire propagation, or a *genesis* in a reverse grassfire growth or via the fitting of primitives such as empty contact disks [50,44, Ch.3 by James Damon]. Alternative field-like structures have been proposed which trade sharpness of details with robustness to noise and small deformations, for example via the medial shape operators of Kovacs [29] or van Tonder [48] *et al*.

I proposed in this short contribution that Information Models (IMs), for example conceived with reference to a linguistic explanatory context and implemented via shape grammars or medial shape operators, could support both perceptual tasks of *genesis* — of meaningful synthetic structures capable of representing perceptual shapes — and *analysis* — of stimuli, producing cues to steer the generative process — seen together as a key building block of the (visual) interface of a mind. Other IMs, for example based on symbolic structures other than derived from natural language descriptions, and of various implementations, should be explored to further test this hypothesis.

Shape percepts are fundamentally based on (learned) meanings the mind constantly projects "out" at the interface that is our visual apparatus, in an act of microgenesis of successive presentations [28]. Without learned meanings (of concepts such as a chair, a bottle, a face, a triangle) there is no shape to perceive and we may remain blind to the nature of objects we are exposed to. When a new scene possibly with unexpected objects in it stimulates our sensors, we can either immediately associate with it some known (shape) meanings or we may create a new set of meanings to become part of our memory banks. An interesting demonstration of this effect can be illustrated for example with the remarkable emergent dalmatian dog percept of R. Gregory and R. C. James [14], in which the dog shape is hardly perceived by first viewers but once it has been cognised it will pop-up even if shown again only years later.

In summary, I started by defining "shape" as a process-based description of an object, which makes explicit useful (to a mind) properties and characteristics of such an object, *e.g.* of a triangle, cube, chair. The Information Model (IM definition) for a shape was proposed as a mechanism to both generate and analyse such a shape percept and situate it at the (visual) interface where a mind can become conscious of it. I also used in my examples some illusory shape percepts — from the Kanizsa family — as these must be explained if a shape theory of human visual perception is to be further studied. Note also that illusory percepts give explicit clues on the potential IMs to put in place [13,18,16]. Much remains to look at, including studying similar 3D shape percepts and extending the family of IMs to cover more applications, *e.g.* in the context of shape grammars.

Acknowledgments

Results in figures 2 ((a) and (b)) and 3 were produced using the shock graph software of Ben Kimia, Ming-Ching Chang, Amir Tamrakar and Nhon Trinh developed at Brown University [50,26]. Results in figure 5 were produced using the BIM platform being developed by the Mutators Research Group, which currently includes W. Latham, S. Todd, P. Todd and myself at Goldsmiths, University of London, and A. Steed at UCL. I thank Andrea Baier and the anonymous reviewers for help in revising the manuscript.

References

1. Andaroodi, E., et al.: Ontology-based shape-grammar schema. J. of cultural heritage 7(4), 312–28 (2006)
2. Arnheim, R.: Art and Visual Perception: A Psychology of the Creative Eye. University of California Press, Berkeley, CA, U.S.A. (1974)
3. Biederman, I.: Recognition-by-Components: A theory of human image understanding. Psychological Review 94, 115–47 (1987)
4. Biederman, I.: Visual object recognition. In: Kosslyn, S.F., Osherson, D.N. (eds.) An Invitation to Cognitive Science, Visual Cognition, vol. 2, chap. 4, pp. 121–65. MIT Press (1995)
5. Bloomenthal, J., Lim, C.: Skeletal methods of shape manipulation. In: IEEE Proc. of the International Conf. on Shape Modeling & Applications. pp. 44–7. Japan (1999)
6. Bloomenthal, J., Wyvill, B. (eds.): Intro. to Implicit Surfaces. Morgan Kaufmann (1997)
7. Blum, H.: A new model of global brain function. Perspectives in Biology and Medicine 10, 381–408 (1967)
8. Blum, H.: Biological shape and visual science. J. of Theoretical Biology 38, 205–87 (1973)
9. Chang, M.C., Leymarie, F.F., Kimia, B.B.: Surface reconstruction from point clouds by transforming the medial scaffold. Comp. Vision & Image Understanding 113(11), 1130–46 (2009)
10. Chomsky, N.: Syntactic Structures. Mouton de Gruyter (1957), 2nd edition published in 2002
11. Deak, P., Reed, C., Rowe, G.: CAD grammars. In: Proc. of Computer Aided Methods in Optimal Design and Operations. pp. 119–128 (2006)
12. Giblin, P.J., Kimia, B.B., Pollitt, A.J.: Transitions of the 3D medial axis. IEEE Trans. on Pattern Analysis & Machine Intelligence 31(5), 900–18 (2009)
13. Gregory, R.L.: Perceptual illusions and brain models. Proceedings of the Royal Society B 171, 279–96 (1968)
14. Gregory, R.L.: The Intelligent Eye. Weidenfeld and Nicolson, London, U.K. (1970)
15. Gregory, R.L.: Images of mind in brain. Word and Image 21(2), 120–3 (2005)
16. Gregory, R.L.: Seeing Through Illusions. Oxford University Press (2009)
17. Grobler, F., et al.: Ontologies and shape grammars. In: Proc. of 3rd International Conf. on Design Computing and Cognition (DCC). pp. 23–40 (2008)
18. Hoffman, D.D.: Visual illusions and perception. In: Yearbook of Science and Technology. McGraw-Hill (2005)
19. Hoffman, D.D.: Sensory experiences as cryptic symbols of a multi-modal user interface. In: Bauer, M., et al. (eds.) Kunst und Kognition, pp. 261–79. Wilhelm Fink (2008)
20. Hoisl, F., Shea, K.: Exploring the integration of spatial grammars and open-source CAD systems. In: Proc. of 17th Int. Conf. on Engineering Design. vol. 6, pp. 427–38 (2009)
21. James, W.: The Principles of Psychology. 1950 reprint, Dover Publications. (Original: 1890)
22. Kanizsa, G.: Quasi-perceptional margins in homogenously stimulated fields. Psicologia 49, 7–30 (1955)
23. Kellman, P., Shipley, T.: A theory of visual interpolation in object perception. Cog. Psych. 23(2), 141–221 (1991)
24. Kelly, M.F., Levine, M.D.: Annular symmetry operators. In: Proc. of the IEEE 5^{th} Int. Conf. on Computer Vision. pp. 1016–21 (1995)
25. Kimia, B.: On the role of medial geometry in human vision. J. of Physiology, Paris 97(2–3), 155–90 (2003)
26. Kimia, B., Tamrakar, A.: The role of propagation and medial geometry in human vision. In: Biologically Motivated Computer Vision, pp. 197–207. LNCS 2525, Springer (2009)
27. Koenderink, J.J.: Solid Shape. A.I. Series, MIT Press (1990)
28. Koenderink, J.J.: Vision and information. In: Perception beyond Inference, chap. 1, pp. 27–57. MIT Press (2011)

29. Kovács, I., Feher, A., Julesz, B.: Medial-point description of shape: A representation for action coding and its psychophysical correlates. Vision Research 38, 2323–33 (1998)
30. Latham, W., et al.: Using DNA to generate 3D organic art forms. In: Applications of Evolutionary Computing, pp. 433–42. No. LNCS 4974, Springer (2008)
31. Lee, Y.T.: Information modeling: From design to implementation. In: Nahavandi, S., Saadat, M. (eds.) Proceedings of the Second World Manufacturing Congress. pp. 315–321 (1999)
32. Lehar, S.: Computational implications of Gestalt theory 2: A directed diffusion to model collinear illusory contour formation (1999)
33. Lehar, S.: The World In Your Head: A Gestalt view of the mechanism of conscious experience. Lawrence Erlbaum Associates (2003)
34. Lehar, S.: The constructive aspect of visual perception (2007)
35. Leymarie, F.F.: Thoughts on shape. In: Albertazzi, L. (ed.) Visual Thought: The Depictive Space of the Mind, Advances in Consciousness Research series, vol. 67, pp. 303–50. John Benjamins (2006)
36. Leymarie, F.F., Kimia, B.B.: The medial scaffold of 3D unorganized point clouds. IEEE Transactions on Pattern Analysis and Machine Intelligence 29(2), 313–30 (2007)
37. Leyton, M.: A process grammar for shape. Artificial Intelligence 34(2), 213–47 (1988)
38. Leyton, M.: A Generative Theory of Shape. LNCS 2145, Springer-Verlag (2001)
39. Lindenmayer, A.: Mathematical models for cellular interactions in development: Parts i and ii. Journal of Theoretical Biology 18, 280–315 (1968)
40. Mueller, P., et al.: Procedural modeling of buildings. In: Proc. of ACM SIGGRAPH. vol. 25, pp. 614–23 (2006)
41. Naeve, A., Eklundh, J.O.: Generalized cylinders - what are they? In: Arcelli, C., et al. (eds.) Aspects of Visual Form Processing. pp. 384–409 (1994)
42. Prusinkiewicz, P., Lindenmayer, A.: The Algorithmic Beauty of Plants. Springer (1999)
43. Sebastian, T.B., Klein, P.N., Kimia, B.B.: Recognition of shapes by editing their shock graphs. IEEE Trans. on Pattern Analysis & Machine Intelligence 26(5), 550–71 (2004)
44. Siddiqi, K., Pizer, S.M. (eds.): Medial Representations: Mathematics, Algorithms and Applications, Computational Imaging and Vision series, vol. 37. Springer (2008)
45. Stiny, G., Gips, J.: Shape grammars & the generative specification of painting & sculpture. In: Freiman, C. (ed.) Information Processing. vol. 71, pp. 1460–5. North Holland (1972)
46. Teboul, O., et al.: Shape grammar parsing via reinforcement learning. In: IEEE CVPR (2011)
47. Todd, S., Latham, W.: Evolutionary Art and Computers. Academic Press (1992)
48. van Tonder, G.J., Ejima, Y.: Flexible computation of shape symmetries within the maximal disk paradigm. IEEE Trans. on Systems, Man & Cybernetics, Part B 33(3), 535–40 (2003)
49. van Tonder, G.J., et al.: Visual structure in Japanese gardens. Journal of the IEICE 86(10), 742–6 (2003)
50. Trinh, N., Kimia, B.: A symmetry-based generative model for shape. In: Proc. of the IEEE 11^{th} Int. Conf. on Computer Vision. pp. 1–8 (2007)
51. Vishwanath, D.: The epistemological status of vision. Axiomathes 15(3), 399–486 (2005)
52. Zhu, S.C., Mumford, D.: A stochastic grammar of images. Computer Graphics and Vision 2(4), 259–362 (2006)