

Carmelo Cali

## **Gestalt Models for Data Decomposition and Functional Architecture in Visual Neuroscience**

### **Introduction**

The claims to bring back Gestalt theory into the Neurosciences research field have been laid on both theoretical and experimental grounds. The constraints upon which things look as they do are commonly held as explanatory and predictive phenomenological conditions that supposedly contribute to a better understanding of brain mechanisms underlying the perceptual world structures. On the other side, some views of Gestalt theory have been claimed to fit with interpretations of structures of receptive fields, and of long-range neuronal connection pathways.

Hence, as theoretical issues unravel, empirical implications arise. Breakthroughs and research programs may amount to evidence that Gestalt theory and contemporary visual Neurosciences cooperate for mutual benefit. But for that to obtain, the following questions are at stake.

Which form is the phenomenological claim of Gestalt theory to be given, and how is it to be implemented in visual Neurosciences? How is one to build a common ground upon which to compare standard and Gestaltist models of neuronal computation? Once presumed that it could provide comparable neuronal models, is Gestalt theory expected to play any role in assessing how much the properties predicted by competing theories of one and the same brain area, say V1, are aligned with the perceptual structures of environment in which the brain evolved?

In order to answer these questions, this theoretical paper defines the characteristics of a standard model that qualifies as a received view in visual Neurosciences. Particular attention is paid to its intended interpretation of visual primitives definition, be it laid down on a somewhat intuitive or an abstract ground. As far as they regard the perceptual system capability of exploiting regularities and structures that can be retrieved in the environment, issues of functional architecture are discussed as well.

Then some critical objections to the received view are presented. The research paradigm of natural images statistics is briefly introduced, which is meant to

***GESTALT THEORY***

© 2013 (ISSN 0170-057 X)

Vol. 35, No.3, 227-264

capture the ecologically valid conditions of perception against which the cortical units' functions and architecture are to be modelled.

Finally, after reviewing the literature on the implications of Gestalt theory for Neurosciences, and of the research into natural scenes statistics that makes empirical reference to Gestalt experimental evidence, the conjecture is made that experimental phenomenology and Gestalt theory can profitably contribute to the research. Arguments are founded on the relevance ascribed to the orderings and structures that carve up phenomenologically the environment for an effectual cognitive and behavioural interaction to obtain.

A formal treatment of modelling phenomenological structures can prove efficient to compare and assess the theoretical and empirical implications for the perception research.

### **1. Varieties of Neuronal Models: Visual Primitives and Functional Architecture**

Since the groundbreaking Hubel & Wiesel (1962), the study of response properties of neurons in cat visual cortex yielded models for the structure of receptive fields (RFs), but also for the function ascribed to visual neurons providing hypotheses for the wiring mechanism underlying brain functional architecture. The RFs are made up of different subunits that respond to luminance increments or decrements, which are spatially organised in different layouts that were deemed to account for the tuning of a neuron's response selectivity to such well-specified stimuli features as position, orientation, length and width. The assumption is that neurons have a structure tuned to local contrast dimensions that reveal locally-oriented structures of objects, and that perceptual primitives are lines or bars. Hence, the neurons are taken as edge *detectors*. Simple cells respond to such stationary stimuli as luminance edges of specific orientation and polarity or as bright/dark lines only if located in appropriate subregions of RFs. Complex cells respond to moving lines and edges in *any* point of RFs with direction preference. Hypercomplex cells showed even more selectivity by responding less to more elongated stimuli than to shorter ones.

The assumption is that neurons at a low level with overlapping and aligned RFs converge to higher level neurons, whose selectivity to preferred stimuli are monotonically constructed in a tuning to more complex features. The spatial arrangement of convergent and divergent projections from low to higher levels defines the connections by which a functional architecture arises that is characterised as a bottom-up feedforward model, since information flows in one direction. Thus the emergence of more complex RFs, tuned to increasing dimensions of optimal stimuli and responding with a quite complex behaviour, is explained.

Hubel & Wiesel (1977) argued that visual cortex is built up as a retinotopic map

organized by an array of neighbouring vertical columns specific for isorientation and ocular dominance. Moving transversally across two such contiguous columns, an orderly cycle of shift in value progression of both parameters is observed. A loop of neurons whose RFs span all these values was called *hypercolumn*. Taking a  $1\text{mm}^3$  column of cortical tissue perpendicularly through V1 layers delivers a cube made up by two hypercolumns of neurons with overlapping RF properties with shifted retinotopic position. Hubel and Wiesel contended that this “ice cube” model is the basic analysis unit of functional architecture. These cortical maps capitalized on an orthogonal layout with each other to optimize their interaction, thus providing «more than enough machinery» to process all possible values of a visual point in space.

De Valois & De Valois (1980) and De Valois, Albrecht & Thorell (1982) laid the ground for a quite different model. The lower level simple cells’ behaviour did not match in every respect the predictions of the Hubel & Wiesel model. These response profiles showed a primary peak sided by small lobes of excitatory and inhibitory activity. Further a variety of RFs size was found whose properties appeared to correlate with different scales of visual spatial information. Finally, when tested with sinusoidal stimuli, simple and complex cells showed different tuning to frequency and orientation.

Then a frequency channel model was developed drawing heavily on psychophysical evidence about contrast sensitivity function (CSF) – that is the relationship between the reciprocal of contrast detection at thresholds and different spatial frequencies – and adaptation with such stimuli as sinusoidal gratings (Campbell & Robson 1968; Blakemore & Campbell 1969). CSF was taken to represent the envelope of many narrowly-tuned frequency channels. It plots the properties of channels sensitive to different ranges of spatial frequency that reach threshold independently. Measurements were then taken of tuning curves with peak sensitivity and orientation bandwidth for each of the independent frequency channels. Since sinusoidal gratings are used as stimuli, a claim to the Fourier theorem was made to take sinusoidal gratings as the visual primitives whose appropriate combination of dimensions – frequency, amplitude, orientation and phase – specify the coarse and fine structure of visual objects and environments. RFs were identified with localized, oriented, bandpass *filters* of spatial frequencies whose power and phase spectra convey the information about the visual structure at different locations and to different scales.

The evidence suggested another functional model (De Valois & De Valois 1988). Cortical layers architecture arranges cells in order according to difference in frequency selectivity perpendicular to orientation preference. Thus each hypercolumn is given a two-dimensional coordinate system where separate neuronal channels decompose locally visual space into a set of primitives. Given

that cells' selectivity appears to fall in a continuum of tuning degrees, neurons map locally the luminance structure of objects, that is at an extent of small sinusoidal patches. The activity of each hypercolumn preserves the general Fourier structure, decomposing it by a series of variables as to location, orientation, scale, and phase at particular retinal positions. Such elements of objects as edges are maintained to be recovered at a higher level by appropriate combination of filter with appropriate form and composition.

**2. The Linear-Non Linear Standard Model**

Given a visual image  $I(x, y)$  as a 2D section of the visual environment, the amplitude  $A$  and the phase  $\psi$  of the frequency  $\omega$  in the  $(x, y)$ -plane takes different values located at any  $n \times n$  point of the image patch. The sinusoidal properties  $A \cos(\omega_x + \omega_y + \psi)$  of the image are also conveniently described in terms of distribution of the coefficients  $C \cos(\omega_{x,y})$  for  $A$  and  $S \sin(\omega_{x,y})$  for  $\psi$ . The weighted sum of the coefficients conveys the information about the essential structure of visual images. Accordingly cortical neurons act as filters for the coefficients of sinusoidal components to extract the "strength" and location of only those frequencies matching the filter's response functions at appropriate scale and location. If the sinusoids are orthogonal, that is the frequency  $\omega_x$  and  $\omega_y$  components in the  $x$  and  $y$  directions are computable as the vector  $\omega$  by the dot product with the coordinate  $(x, y)$  vector, then the cosine and sine functions provide the image basis whose projections give the coefficients for many different frequencies:

$$\sum_x \sum_y I(x, y) \cos(\omega_x, \omega_y), \sum_x \sum_y I(x, y) \sin(\omega_x, \omega_y)$$

Information about the different frequencies is deemed to be processed independently in spatial frequency channels made up by neurons according to RF's response property selectivity and limited extension: e.g., for  $A$  a neuron convolutes a spatially delimited area with the product of a weighting and a window function:

$$\sum_x \sum_y I(x, y) W(x - x_0, y - y_0) \cos(\omega_x(x - x_0) + \omega_y(y - y_0))$$

where  $W(x, y)$  is a weight function that spans the area delimited by a window centred at  $I(x_0, y_0)$ , and the output  $A(x_0, y_0)$  is dependent upon the amplitude response function of the neuron. That convolution obtains an amplitude and phase map of the frequency for every localized area selected by the window function. Drawing on Marceljia's (1980) mathematical model of cortical RFs, Daugman (1980) hypothesized that RFs had similar properties to Gabor linear filters, that is filters whose components are in phase and are modulated by a Gaussian window.

Originally designed for one-dimensional analysis of acoustic signals, Gabor filters were extended to two dimensions to apply to the visual system (Daugman 1985) analysis in space and frequency domains. It is worth noting that, theoretically, the typical output of Gabor filters resembles the response profile with the additional lobe of activity found by De Valois & De Valois (1980). Jones & Palmer (1987) tested the optimality of Gabor function to fit the cat visual cortical neurons' properties. Daugman (1988) proposed a neural network model for Gabor units' analysis, segmentation and compression of images. Gabor linear filter function was assumed to provide a meaningful mathematical expression of the form of simple cell RFs. The form of a complex Gabor filter is the following:

$$g(x,y) = s(x,y) w_r(x,y)$$

where  $s(x,y)$  is the complex sinusoidal carrier that specifies the spatial frequency and phase in cartesian or polar coordinates;  $w_r(x,y)$  is an elliptic gaussian envelope of a determinate magnitude that specifies the location of peak response function, and the coefficients scale along the  $(x,y)$ -axis, along with a parameter for rotation or translation.

The critical tuning properties of a 2D Gabor filter for the visual system for spatial frequency, orientation and phase angle are constrained in width and length by the gaussian modulation of variance in the  $(x,y)$ -directions, with the centre location at arbitrary spatial coordinates obtained by translation, and arbitrary orientation obtained by rotation anticlockwise. The output of a neuron that behaves as a Gabor filter is the extraction of local structure from a delimited image patch with tuning degree that decreases at increasing distance from the centre of RF.

Field & Tolhurst (1986) analysed data from a variety of simple cells RF whose response profiles as to bandwidth and symmetry were best fitted to the size and shape of Gabor functions.

Results were extended to wavelet analysis, that is Gabor functions whose gaussian envelope is a constant number of cycles of the sinusoids. Field (1993) meant to understand the information which power and phase spectra convey about the visual structure, how local phase spectrum is aligned along neighbouring frequency bands, and how power and phase invariance are preserved across scales. The filter function model was developed in its functional architecture implications. The image structure is mapped by cells arranged in cortical layers as a multiscale pyramid of local oriented filters at different spatial frequencies such that for each layer orientation tuning is represented along  $(x,y)$ -positions for each range of high, medium and low frequency tuning. At isorientation, cells are arranged in different positions and with wider bandwidth at increasing frequency. Each layer maps a spatial sampling grid for 2D filters, whose sampling distance

is proportional to RFs size, which is determined by the selected basis function that extracts the inherent structure of image components. If the basis functions are wavelets, then that arrangement of filters may decompose visual structures in power and phase spectra by a set of shifted, rotated and dilated copies of the same basis functions located at different layers.

In fact, these models have been reconciled in abstract modelling within the theoretical framework provided by the *linear system analysis* in order to accommodate the enriched view of the functional architecture as well (Marmarelis & Marmarelis, 1978).

Stimulus selectivity is no more analysed in terms of intuitive visual features, say edge, by which their function is to be characterized, say detector. Each neuron-type function is conceptualized in a neutral way as a filter operation. The determinate range of values to which neurons specifically respond is represented as a complex domain of a multidimensional stimulus space. Features are the unique subsets that neurons filter out from multiple dimensions. Selectivity is expressed by mathematical equations that describe the specific tuning properties of neurons. The loss of intuitiveness of feature identification is compensated by a gain in abstract quantification that ensures a high level of confidence in modelling responsiveness to such diverse stimuli as contrast, spatial frequency, orientation, position, motion, colour and time (Albrecht, Geisler & Crane 2004)

Indeed, the simple cells in the Hubel & Wiesel model meet the response properties of the linear system, e.g. superposition and additivity. They show:

1. distinction of RF subregions;
2. spatial summation within a given subregion;
3. mutual antagonism between subregions;
4. predictability of responses to new arbitrary stimuli by spatio-temporal summation for (1-3);
5. composition of response to more complex stimuli by a weighted sum of responses to simpler ones; (Hubel & Wiesel 1959; Movshon et al. 1978a,b; Albrecht et al. 2003; Carandini et al., 2005)

A visual neuron is deemed to perform the following linear filter processing

$$r_j = \sum_{x,y} W_j(x,y)I(x,y) + r_0$$

that is a weighted sum of image properties, with the response output  $r_j$  above the spontaneous firing rate ( $r_0$ ), where the weighting function  $W_j(x,y)$  describes the RF as it is mapped by reverse correlation with such different stimuli as a luminance spot flashed on different points of a visual field, edges, lines or bars, or sinusoidal gratings. Linear system identification models neurons as linear filter

operators in a way that is alleged to be independent from any assumption about primitives or theoretical principles, since it tests how a response is affected by a stimulus dimension component as is the case in regression analysis.

Hence the stimuli features selectivity and response optimality of neurons are characterized with the known properties of linear filter operators by mapping the neural responses in the stimulus space-time:

$$r_i(t) = \iiint x, y, \tau W_j(x, y, \tau) I(x, y, t - \tau) d\tau, dx, dy$$

(adapted from De Angelis & Anzai 2004).

But essential characteristics of neuronal response and non-linearities are missing, which have been observed since the first experimental works to model linearly and measure neuronal response profiles.

Complex cells do not meet linearity conditions. But some non-linearities were observed in simple cells too.

Though quantitative tests measured the responses of simple cells to single optimal stimuli as linear summation outputs, the linear model does not predict correctly orientation and frequency tuning magnitude response to non optimal stimuli (De Angelis, Ozhawa & Freeman 1993). The linear model allows for firing rates smaller than  $r_0$  to give a negative contribution to the linear sum, that which predicts response output being negative due to the fact that  $r_0$  tends to be small. Were their response magnitudes highly maintained, neurons could simply respond less than a spontaneous rate and have negative values. But unlike the linear filter that is not the case. In Heeger (1992) a half-wave rectification mechanism is proposed by which two complementary cells with reverse polarities respond such that the weighting function  $W(x,y)$  results from positive ( $W_i$ ) and negative ( $-W_j$ ) weights. Respective outputs replace each other and one of the two at a time gives a nonzero response.

Along with complex cells, simple ones show response saturation at high contrast (Albrecht & Hamilton 1982; Maffei & Fiorentini 1973), whereas a linear operator's response should increase. Stimulated with a superimposed pair of gratings with orthogonal orientation and direction, simple cells reduce their response unlike linear operators whose output should not be affected for non-preferred stimuli (Bonds 1989; Albrecht & Geisler 1991). Other cases of "non specific suppression" from stimuli whose presentation alone brings about no overt response have been tested (De Angelis et al. 1992). Therefore a normalization model has been put forward to account for suppression: each neuronal linear response is proportionally divided by the pooled activity of the responses of neighbouring neurons, against which it is rescaled (Carandini & Heeger 1994;

Heeger, Simoncelli & Movshon 1996). Since divisive suppression increases with stimulus contrast and involves pooled cells with different orientation tunings, that model explains response amplitude saturation and cross-orientation inhibition. The amplitude response  $r_j$  of a neuron with a RFs weighting function  $W_j(x,y)$  that integrates over contrast  $C(x-x_0,y-y_0)$  in an  $I(x,y)$  image patch is scaled accordingly to

$$r_j = \left[ A(x, y) \frac{c}{\sqrt{\sigma(f)^2 + c^2}} \right]^n$$

where  $c$  is contrast,  $\sigma(f)^2$  is determined by normalization and grows with temporal frequency  $f$ , and  $n$  is the exponent of rectification. In the early visual system another important source of non-linearity is due to light adaptation. Since the dynamic range of neurons in LGN and V1 to the light intensity that is locally prevalent in the visual field is limited, their weighting function is affected by drastic luminance and contrast changes in natural viewing conditions (Heeger 1992; Schwartz & Simoncelli 2001). Control mechanisms are required that modulate the gain and the integration time of neuronal response, that is the height and the width of  $W(x,y)$  that filters different delimited areas of an image patch at different times with root-mean squared luminance and contrast standard deviations (Shapley & Victor 1978; Zhongloul et al. 2005). Then, luminance and contrast variances are supposed to be divided by the norm of the image patch

$$c \sqrt{\sum_{x,y} I(x, y)^2}$$

which may occur as denominator for linear weight summation in a divisive normalization model.

Complex cells are sensibly non-linear because of stimulus polarity and phase invariance that cause them to mix ON and OFF responses throughout the RF. That makes averaging triggered spikes by reverse correlation analysis ineffective for measuring their RFs and quantitatively estimating the spatio-temporal functions of their subfield components. Instead, a covariance matrix is computed from the spike triggered response set to get those eigenvalues that belong to significant eigenvectors of the stimulus set that induced cell responses. Eigenvectors are assumed to be the features of the stimulus space that drive the cell activity, whose properties match the response profile of the units of a complex cell against which RFs are mapped.

We find that spatial profiles of eigenvectors for each complex cell resemble Gabor



functions with similar spatial frequency tuning which differ only by a phase shift of approximately  $90^\circ$ , and that the contrast response function of a complex cell is captured by the so-called energy model

$$r_j = \left( \sum_{x,y} W_{j1}(x, y) I(x, y) \right)^2 + \left( \sum_{x,y} W_{j2}(x, y) I(x, y) \right)^2$$

The energy model captures the idea of non-linearity due to second order interaction that adds to the first order convolution filtering of (4). In case of responses driven by the combination of two or more stimuli the spatio-temporal RF map embeds a non-linear map:

$$r_k(t) = \iiint x, y, \tau W_f(x_j, y_j, \tau_j) I_1(x, y, t - \tau_j) d\tau_j dx_j, dy_j + \\ + \iiint x, y, \tau W_{NL1}(x_{j1}, y_{j1}, \tau_k) I_2(x, y, t - \tau_k) d\tau_k dx_{j1}, dy_{j1}$$

where  $W_{NL}$  stands for the first non-linearity, that is the second order interaction map that weights two stimuli  $I_1$  and  $I_2$  (adapted from De Angelis & Anzai 2004). A more comprehensive model that embeds simple and complex cells' non-linearities is proposed to account for modulation of responses due to contextual effects and neural response history (Heeger, Simoncelli & Movshon 1996; Carandini, Heeger & Movshon 1997; Carandini et al. 2005).

### 3. The Standard Model Questioned: Theoretical and Anatomical-Functional Issues

Though the authors contributed greatly to the definition of the linear-non linear model, Olshausen & Field (2006) argue boldly that most of its descriptions of early vision are highly questionable.

They contend that the role ascribed to non-linearities and interactions of responses is not sufficient since its elementary and reductionist approach, whose assumptions lead to biased choices of sampling techniques, stimulus set, and theories in such a way that the percentage of V1 function that is not understood is estimated to be 85%. They cite laboratory findings that maintain that models exhibit a very low level of prediction, in particular if applied to ecological conditions, thus leaving much of the observed variance unexplained.

As far as the experimental stimulus set is concerned, sinusoidal gratings are not as fit to study vision in ecological conditions, where few things oscillate spatially or temporally, as it might appear on theoretical and mathematical grounds. Further, they are meaningful stimuli for linear space-time invariant systems,

and the Fourier coefficients provide only one of the possible basis functions to probe the visual system, that is also likely to be unfit for highly non-linear units. Light edges, bars, line terminations, Gabor functions and Walsh patterns are effective at best fitting the models. But they perform poorly at predicting actual neurons' responses in natural viewing conditions (David, Vinje & Gallant 2004 who compared predictions of linear weighted sum and a phase-separated Fourier models for non-linear complex cells).

White noise, that is stimuli with a flat power spectrum at different frequencies, and M-sequences, binary nearly white stimuli (Reid, Victor & Shapley 1997), presented in sparse or dense noise conditions, can afford a wider range of stimuli such that some specific forms of non-linearity are mapped out. But the confidence in a complete mapping of non-linearity is hindered by the huge combinatorial explosion of dimensions in experimental design of conditions for mapping a highly non-linear stimulus space.

Theoretical assumptions are not the least of biases. Theories about visual system functions seem to strike a sub-optimal balance in trading the questions about the problems vision is bound to solve off against the explanation of a particular subset of data. The authors mention the Fourier view of vision as a good case in point. Attestation comes from Wilson & Wilkinson (2004): the fact that both spatial frequency and orientation bandwidth decrease with increasing peak frequency indicates that spatial processing cannot be accurately described by either a Fourier or a wavelet transform.

Furthermore, whether the categories of simple *vs.* complex or hypercomplex cells and the parameters to measure bi-modal distribution are attested by actual neuronal responses or derive from the way neurons were stimulated is disputed. Sound parameters for estimating distances, spacing and overlaps among units in the neural space are useful for evaluating theories. According to Olshausen and Field, the energy model does not work well for responses of cells with a partial overlap of activating zones.

All those arguments are brought to bear in identifying the visual primitives from which theories or experimental probes of visual system are defined. The hypercomplex end-stopped cells seemed to imply models that decompose visual structure as a "drawing-like blocks world" made up of bars and line terminations. But inspection of a scene occurring in natural conditions does not reveal many instances of such units while structures appear to be more complex than that even at the scale of V1 RFs. When theories assume that cells' filtering activity for detecting luminance discontinuity subserves coding for such features as contours, junctions and corners upon which to build edges of visual objects, models and functional questions risk being misled by dealing with such post hoc alleged primitives. Eventually, were images in natural conditions composed of such

features, a crowded multifarious collection of orientation signals would ensue that would be hard to make sense of for linear operators merely data-driven.

Finally, the variance of V1 neuronal responses in ecological conditions for which the standard linear-non linear model is able to account is a strikingly small proportion, estimated as 30%-40% by David, Vinje & Gallant (2004).

Matters become more unsatisfactory if neuronal interdependence and contextual modulation of responses are at stake. Neuron responses appear to be modulated by stimuli at different degrees of eccentricity far from their RFs (Maffei & Fiorentini 1976). Gilbert & Wiesel (1990), Kapadia et al. (2000) found that flanking elements in outer surrounds modify a neuron's response to optimal stimuli within its RF, even though the presentation of the former in isolation does not evoke any response. Cross or iso-oriented stimuli may enhance or suppress the response-inducing contextual effects that seem involved in surface perception, figure ground segregation and border ownership (Lamme 2004). All this hints at non-local integration among RFs, and the bottom-up feedforward constraint for the functional architecture is deemed to have been deeply reconsidered (but see Reid & Usrey 2004). Along with feedforward connections, long-range lateral horizontal projections between areas on the same layer and feedback reentrant projections from higher levels are presumed to rule modulation and integration across neuronal responses beyond the classical RFs (Gilbert 1992, 1993). Fitzpatrick (2000) reports a complex picture of the arrangement and functional connections of neural projections. On the one hand, the contextual modulation of a neuron's response to an optimal stimulus within RF is found to depend on the contrast magnitude. Low contrast flanking co-linear stimuli facilitate neuron response to optimal stimuli, while high contrast ones inhibit it. This may be a demonstration that length summation areas of RFs are not fixed, but vary as a function of contrast. On the other, discontinuities in the topographic mapping of orientation selectivity were observed such that cells without overlapping RFs fall in close proximity mediating orthogonal oriented contextual effects. Local horizontal connections seem to provide common inputs for cells with different RF properties. Furthermore, iso-orientation inhibitory flanks beyond the classical RFs are reported to be asymmetrically distributed.

Therefore, the functional effects of short and long-range connections seem to be so intermingled that the very distinction between vertical and horizontal dimensions of connectivity appears an oversimplification.

On the same evidence, Olshausen and Field argue that the reductionist approach of linear-non linear models to non-linear interactions cannot account for that complex anatomical arrangement and functional architecture, since their piecemeal approach leads to the attempt at accommodating them with additional presentation of surrounding stimuli. But that approach faces combinatorial explosion.

The authors suggest that new theories are needed, guided by functional considerations, which can account for what the natural viewing conditions demand that the visual system undertake effectively: such tasks as contour integration, which requires selective connections within and between V1 layers beyond the scope of the standard model, and the detection and parsing of 3D structures. Then, they suggest that the specification that interaction of information at a local and at a global contextual level, of functional connections of subunits within the RFs and interaction beyond, may be acquired neuronal activity should be tested with the structure of natural images.

#### 4. Models of Neuronal Functional Principles with Natural Image Statistics

In fact, the tenet that questions concerning the visual system's functional principles and its criteria of optimality and efficiency must be addressed through the study of the structure and regularities of natural environment was put forward already by Attneave (1954) and Barlow (1961). Since the visual system must solve the problem of recovering essential properties of the environment, it is likely that its design and efficiency are highly constrained by the environment structures for developmental and evolutionary reasons. Hence, visual system design and strategy may be effectively studied when tested with so-called natural images, that is an ecologically valid selection of stimuli from natural viewing conditions.

In the set of equiprobable randomly-generated images, natural images form a small subset characterized by high redundancy, while for small image patches (8x8 pixel) only a small fraction of white noise images is estimated to give back the regularity of natural images (Chandler & Field 2007).

Regularities in natural images come in many forms at different statistical levels. At the level of one-pixel statistics, light intensity variables are densely distributed around small values. For a system with fixed response range, an efficient strategy is adapting that background luminance level and transforming the raw intensity values into a more uniform distribution, thus maximizing the information capacity.

The log function

$$\phi(x, y) = \ln[I(x, y)/I_0]$$

where  $\phi(x, y)$  is the log-contrast,  $I(x, y)$  the intensity value at a particular location, and  $I_0$  the intensity value such that

$$\sum_{x,y} \phi(x, y) = 0$$

represents a way by which the system enhances contrast response detecting the log-contrast fluctuations against a mean level (Ruderman 1994).

At the level of second order statistics, images with different contents selected from the real world do not show a flat amplitude distribution. Instead, the physical energy peaks at larger scales and decreases at smaller scales, that is amplitude spectra of natural images are stronger at lower than at higher spatial frequencies. Amplitude spectra computed at different frequencies and averaged for different orientations have a  $1/f$  shape, since they have a peak at the lowest frequencies that falls off with increasing frequency. Accordingly, power spectra are approximated by  $1/f^2$ , that is the squares of amplitudes in natural images fall off in a way inversely proportional to the squares of the frequency. This characteristic has been taken to express the scale invariance of structures in natural images. The invariance means a high correlation of pixel values that is stronger for lower spatial frequency components. The fall off exponent has been found to depend upon different categories of natural images. The amplitude spectrum is also linked to depth: at smaller distances it is more isotropic in orientation and contains high spatial frequencies; at greater distances it is stronger along the horizontal/vertical axes and contains lower spatial frequencies.

At the level of higher order statistics, phase spectrum is assumed to be of momentous importance for the perception of image structures of the natural world since its disruption via different degrees of randomization appears to impair perceptual recognition capabilities. Such regularities as relative cross-scale spatial phase alignment preserve the visual saliency of image structures (Hansen & Hess 2007).

Given that those forms of regularity are characteristic properties of the subset of natural images, it was argued that an effective way of studying the visual system was to analyze the data structure and to search for RFs properties and neuronal functional principles that match the structure. That way, the research would not be biased by a priori assumptions about what features should be like, which visual primitives were taken to be mapped, or hypotheses about the nature of vision. Many theories covering computational principles and implementation questions were proposed to match analysis of natural image statistics and provide an explanation of why V1 RFs have a particular shape.

Bell & Sejnowski (1997), van Hateren & Ruderman (1998), van Hateren & van der Schaaf (1998) worked out an independent component analysis (ICA) model of how the natural stimuli statistics trigger neuronal responses and why the shapes of neurons tuning profiles are generated. Field (1993, 1994), Olshausen & Field (1996; 1997) proposed the alternative model of a sparse code (SC) with an overcomplete basis set. Both models stem from Barlow's (1961) insights. Faced with the huge correlations in sensory data, a visual system must aim to

reduce redundancy in order to extract essential information in terms of statistical independent units, which would correspond to the object-causes of sensory phenomena. The maximization of information ensues from storing only priors instead of joint probabilities about the environment, and employing efficiently each neural unit in transmission without any duplication.

For instance, Field (1987), Atick & Redlich (1992) reasoned that images' second order statistics would lead the visual system at an earlier stage to decorrelate responses by whitening visual signals. A linear rise in sensitivity at increasing spatial frequency, inversely related to  $1/f$  amplitude spectrum until it falls off when signal power becomes less or equal to noise, eliminates the correlations, and the shape of such a filter was found to resemble the retinal and LGN centre-surround RFs.

Hence the visual system is assumed to model the distribution of states of environment that cause vision and preserve only those coefficients of structures that match the actual distribution of real world data. A model of an image from the real world is commonly understood as a linear superposition of features:

$$I(x, y) = \sum_i a_i \phi(x_i, y_i) + v(x, y)$$

Where  $\phi_i$  are the basis functions, that is the spatial features that describe the image structure,  $a_i$  are the coefficients that capture the extent to which the features are present in the image,  $v$  is a gaussian quantity regarding all the independent components differing from  $\phi_i$ .

ICA and SC predict the RFs shapes by computing  $a_i$ , that is defining a set of filters whose values represent how image structures are described by a determinate basis set. The values of coefficients and V1 RFs are compared directly or via a simulated reverse correlation analysis with instances of natural images. Thus predictions may be tested by evaluating the distribution of tuning properties. Filter coefficients and estimated RFs show resemblance in terms of being bandpass, oriented, localized and spatial frequency-specific (though see Ringach 2002).

But evidence about V1 attests to an expansion of number of neurons and connections compared with earlier pathway stages. Hence an increase rather than a reduction of redundancy obtains. Barlow (2001) rephrased the goal of the visual system as modelling redundancy to obtain a meaningful representation in such a way that correlation among some units are to be allowed. Since the actual distribution of data drawn from natural images is non-gaussian, but rather satisfies a kurtosis distribution, theoretical and experimental attempts were made to prove that the underlying structure of natural images is coupled with a sparse code and an overcomplete basis set.

Sparseness is defined as the distribution of coefficients whose values are mainly concentrated around zero and have long tails. It applies to neuronal response selectivity, RFs spatial frequency bandwidth and aspect ratio (length/width). When it is an attribute of a distributed code, sparseness allows for punctuated activity of single units that convey more meaning about the structure features to which they are tuned (see Foldiak 1995 for an opposition of sparse models to “grandmother cell” models). Further, sparse activity of few neurons lowers the probability of false matching in pattern recognition.

Overcompleteness defines a code which does not sample critically the input space, that is whose basis functions exceed input dimensionality. To adapt basis function in order to maximize the data likelihood, the visual system is to compute the set of coefficients which are maximally sparse with an overcomplete basis set.

That way, it would bring about an optimal choice of basis functions to map the structure in natural images (through sparseness), and at the same time an efficient strategy in tiling the joint space of multidimensional stimuli parameters to match how many of them are instanced in the structure at a particular point.

Different strategies are being proposed for achieving sparseness with such a code. But overcompleteness forces all to face the problem of infinite possible values of  $a_i$  to solve equation (12). In terms of a mean squared error, there are many ways to explain image structures with different families of  $\phi_i$ . Accordingly, it is presumed that the maximization of image data likelihood by the visual system takes the shape of a bayesian inference. Assumed priors over how data images are generated and the distribution of its causes, the solution consists in determining the probability for each basis set of equation (12) given the chosen coefficients:

$$P(\mathbf{a}|\mathbf{I}, \theta) \propto P(\mathbf{I}|\mathbf{a}, \theta) P(\mathbf{a}|\theta)$$

where  $P(\mathbf{I}|\mathbf{a}, \theta)$  rates the probability of an image structure given determinate coefficients  $a_i$ ,  $\theta$  stands for such parameters as  $\phi_i$  and noise variance,  $P(\mathbf{a}|\theta)$  rates the probability of  $a_i$  for some determinate parameters and the distribution induced by a chosen prior function. All this would give the relative probability of different explanations for  $I(x,y)$  structure.

Then the specification of which particular solutions for  $a_i$  to choose is searched for by *sparsification*, a non-linear transform operation that favours emergence of highly localized and punctuately activated units that selects only  $\phi_i$  which best match  $I(x,y)$  and inhibits all other competing bases. In a neural circuit, that means that by leaving  $\phi_i$  unparameterized and having them adapt to natural images, a Hebbian learning rule over  $a_i$  will select adapted  $\phi_i$  that converge to optimal RFs (Olshausen 2004). Each coefficient value is achieved not only by



the feed-forward inner product between basis functions and images, but also by a weighted inner product between the activations of overlapping coefficients, and a cost function ( $-S a_j$ ) that self-inhibits the activity of each unit's activity that is so driven to be sparse.

Vinje & Gallant (2000; 2002) found that the increase in contextual information in natural images induces sparser activity among neuronal units than expected with a simple convolution model. When there is little context dependency in images, there is more ambiguity in choosing the best fit basis functions and the responses of units are approximated by a linear-non linear model. Instead, with highly contextual images the activity of neurons is driven by competition in finding the best functions at any point for a given quantity of structure at a time with a combination of suppression and selective enhancement.

This implies more complex non-linearities than those that the standard model of pointwise non-linearities, added to an earlier linear weighting summation stage, is able to account for. The activity of simple cells involves interaction among units with overlapping RFs, and complex cells' non-linear selectivity seems to require models of coefficient grouping that extract invariance from overlapping subunit pools into non-overlapping RFs.

Hyvärinen, Hurri & Hoyer (2009) argued that neither a combination of features from a linear vector subspace nor an interpolation that takes basis vectors averaged from nearby templates are effective in learning RFs invariant to position and phase. Then they proposed a sparsified energy model with higher level units that squares the sums of underlying pooled outputs but with the proviso that dependencies of lower subspace units are preserved in the form of energy correlations in order to induce a topographical organization of the arrangements of units that resemble the visual cortex.

This points to the main question that natural image statistics poses to the vision research community: the forms of higher order statistical dependencies. That seems to be related to global properties of structure that are likely to be linked to the functional meaning of the so-called beyond classical RFs effects and to the horizontal connections.

Schwartz & Simoncelli (2001) argued that horizontal connections are designed to remove the magnitude correlations via divisive normalization in such a way as to preserve contextual effects in V1 cells. But Olshausen (1997) and Olshausen & Field (2006) proposed incorporating dependencies in sparsified models that do not achieve complete independence of components without committing to the shortcomings of correlations in a code with non-orthogonal basis functions. The redundancy needed to capture more global properties of structures via long-range interactions is preserved in spite of that due to simple RFs overlaps.



## 5. Gestalt Neuroscience and the Natural Image Statistics Paradigm

On theoretical and experimental grounds, many claims were made that phenomenological principles and lawlike grouping factors discovered by Gestalt Psychology could constitute an effective probe for brain mechanism search particularly when non-additive or linearly independent effects are involved.

Jung & Spillmann (1970) qualified as *perceptive field* the psychophysically specified perceptual counterpart of a RF. Spillmann & Ehrenstein (1996) defined the local RF structure as a functional “micro-Gestalt”. The perceptual field, tested with say Hermann grids, was found to resemble noteworthy structural properties of RFs (Ehrenstein 2001). Perceptual equivalence classes for neural mechanisms were identified through psychophysical methods by testing with conditions where interaction or contextual effects defined the very perceptual judgments such as in transposition or frame of reference perceptual judgment tasks (Sarris 2006).

The effects of non-local and long-range connections are reviewed by Spillmann & Werner (1996) and Spillmann & Ehrenstein (2004) who claimed that definition of Gestalt rules making things look as they do helps in specifying basic visual neurons functions. That is the case with grouping rules, figure ground segregation, contour integration, perceptual invariances, contextual effects in abutting gratings, boundary and surface belongingness and motion perception. Indeed, Spillmann (1997; 1999) emphasized that neuroscientific findings that followed Gestalt-like hypotheses marked an advancement in understanding perceptual brain mechanisms.

Finally, Ehrenstein, Spillmann & Sarris (2003) contended that what Gestalt psychologists discovered serves as a phenomenal probe for the search of brain mechanisms and architectural functions.

A case for how phenomenological Gestalt evidence may serve the assessment of research assumptions is made by Wilson (1999) who takes advantage of Kaniza’s abutting gratings to show the shortcomings of Fourier basis functions to meet perceptual behaviour in such cases, and to argue the need of specifying non-Fourier processing for visual neuron computations.

Furthermore, Gestalt psychology’s theoretical assumptions and experimental findings can be applied to some questions raised by the research in the natural scene images paradigm and to their implications for the visual Neurosciences along the lines laid bare by Olshausen & Field (2006).

Hyvärinen (2010) points out that the statistical regularities in natural images represent ecologically valid conditions to model the RFs of simple cells, the emergence of complex cells’ RFs and the occurrence of Gestalt effects. Since Barlow (1961) and Field (1994) the adaptation to the environment properties was made the key question of the research into the neuronal basis of the

perceptual system. Actually, Brunswik & Kamiya (1953) and Brunswik (1956) maintained that Gestalt grouping rules are tuned to statistical regularities of the environment. And most subsequent research was devoted to working out a probabilistic interpretation of Gestalt grouping rules according to geometrical and computational models, which could stand up against the empirical findings on the structure embedded in natural scenes, and at the same time equate human perceptual performances.

Elder & Zucker (1996) developed a Bayesian model and a sparsely-connected graph for a contour perceptual grouping algorithm based on a strong global constraint for closed bounding contours in natural images. Its efficiency improvement in rendering perception stems from computing contour as a global feature, once a saliency measure is provided, in a way that is not reduced to the summation of local cues. This accounts for such unavoidable properties of natural scenes as occlusions, and for the distinction between an open extended chain of edges and bounding contours. On the ground of Gestalt principles of perceptual organization, Elder & Zucker (1993; 1998) tested a geometrical specification of perceptual closure, contrasted with topological closure, as a primitive feature of visual scenes in shape discrimination tasks in order to match the perceptual system's effectual ability to integrate or separate fragments of objects when occlusions, shadows, edges or low contrast occur.

Elder & Krupnik (2001) require object knowledge in the study of how grouping allows objects to appear. They propose a general framework where conditions for regularity of objects, such as simplicity, closure and completeness, are embedded in priors for the estimation of boundaries by a computational model that combines local and global cues. Elder & Goldberg (2002) make use of the insight of Brunswik's experimental design to have an observer trace contours in natural images in order to investigate the respective inferential power of *ceteris paribus* factors such as proximity, similarity and good continuation, in order to define quantitative probabilistic analysis of the use that the perceptual system makes of information about depth and segmentation of contour occlusion.

Sigman et al. (2001) and Geisler et al. (2001) found that geometrical relations in pair-wise statistics of edge elements show local maxima in their probability distribution when they are approximately at a tangent to common circles and to parallel lines. They argue that these values of relative orientations of neighbouring segments match expectations for what natural image structures would look like were Gestalt grouping rules of good continuation proved to be sound in extracting smooth contours in occlusion cases. That is suggested to provide further evidence for anatomical-physiological findings on long-range facilitatory projections among pyramidal neurons with iso-orientation preference, and psychophysical measures of contour grouping based on alignment, orientation and axial offset.

In fact, it can be sensibly held that Gestalt theory does not abide by a central tenet of the research based on natural images paradigm. Actually Olshausen (2004) claims that the key question for Neurosciences, that is “how neurons respond to a stimulus”, cannot be answered before the more fundamental one is settled: what it is for something to be a stimulus in a natural scene. But the latter is an ill-posed problem that needs to be specified further with all the available contextual and high level knowledge to be ascribed to the perceptual system. Accordingly, the only explanatory strategy is modelling natural images to study how neurons respond assuming that priors are constructed upon their statistics that the visual system uses to infer world properties from incoming data, and under the hypothesis that these priors are embedded in the neuronal circuitries. But Gestalt theory principles would be strongly at odds with describing the visual system problem solving as the result of a mere adaptation to the structure of the visual environment by an inferential process that defines priors over the physical causes of perception and then adapts through a maximization of posterior distribution. Since their polemic with the theory of perception put forward by von Helmholtz (1867/1962), Gestalt psychologists claimed that the recourse to inferential models is often justified by a poor acquisition of data that does not recognize the phenomenological structures that build up the order of the perceptual environment.

But contending that Gestalt grouping rules reflect statistical properties of the environment does not need to commit to a cognitively inferential theory of perception that reduces perception to detection of local features in the stimulus array of the visual scene devoid of any order or structure whatsoever. Instead, that would amount only to suggesting that Gestalt rules are tuned to regularities in the way environmental objects appear, whose statistical properties overlap the phenomenological conditions discovered by Gestalt psychologists. This point seems consistent with Koffka (1915) who shows how the question of what it is for something to be an environmental stimulus for the perceptual system is settled with a phenomenological specification that takes into account the local underdetermination of being a stimulus that requires reference to contextual information.

On the other hand, it is true that natural images are employed to derive a state space for the receptor array. At each dimension, every point specifies the complete set of receptor values for any image. But that does not imply that their use is limited to such a construal. Geisler (2008) claims that more relevant statistics are obtained if measures are not completely confined within the domain of images. Rather measurements both within images and the corresponding environments allow to obtain “ground truth information”, that is the use observers make of image structures when they interact with the environment (Wilson, Geisler & Perry 2009 use “across-domain” statistics to extract edge elements’ geometry and

contrast polarity from natural images, and derive a model for an ideal observer performance in occlusion tasks). Hence, a principled approach to natural image scenes analysis takes Gestalt grouping rules as an effectual means of retrieving the order in which to integrate elements of natural environment in such a consistent way as to get ordered appearances of articulated objects through their high level properties (i.e. curvature, closure, concavity and convexity), instead of a mere aggregate of features. And that is the main tenet of theoretical and experimental Gestalt psychology, which allows for a construal of natural images in terms of the relevant subset of phenomenological conditions.

Mahamud, Thornber & Williams (1999) developed a saliency measure for closure as a perceptual Gestalt property that can be used to detect bounding contours of unknown objects in natural images. Grouping is delivered through segmentation without a-priori knowledge, rather imposing a subset of Gestalt principles for perceptual organization. Such a constraint for proximity and smooth continuation clues is meant to allow for the segmentation of strongly connected units, and not only edge chains or open contours, to obtain.

Accordingly, a suggestion can be made that natural scene analysis highlights structures whose construal is two-fold: as regularities displayed by images at various dimensions and specified in the state space of a receptor array, and as correlates of phenomenological conditions of orderings that make things look as they do for the perceptual system of organisms that behave rationally and purposively in the environment.

Some theoretical implications would ensue. Olshausen (2004) emphasizes that a principled approach to discovering the way things are grouped together is needed in order to understand how the visual cortex uses image structures to capture what is “out there”. For a meaningful representation of sensory information to obtain, SC models suppose that the world can be described in terms of a small number of responses by a small fraction of active neurons in the population at any given moment, but in such a way that the cortex must also choose how to tile the entire joint space of relevant parameters to get a complete representation of the image. Assuming distributed sparseness and overcompleteness, SC models trade off the meaningfulness of the activation of a small number of units for each parameter against their combination, that which allows a convenient form of world description to obtain in that each output would recover the amount of structure occurring at determinate image locations in terms of features specifically tailored to structures occurring in natural scenes. This is much like the form of language that allows the generation of complete descriptions by sentences that combine words, which are more than phonemes, while any given sentence utterance realizes only a small fraction of the available words.

Gestalt psychology made quite the same point when applied to the way objects

appear in the environment. Grouping laws keep a balance between having few enough rules and conveying efficiently the meaning of environmental object structures. For there are many more perceptual structures than grouping rules, and on the other hand any given perceptual unit is an instance of only a subset of the grouping rules that hold competitive and cooperative forms of interaction with one another.

Therefore as far as the models of visual cortical units and processing are concerned, Gestalt theory can provide a rationale for the hypothesized RFs and cortical architecture. The outputs of neuronal units should be both sufficient to recover the structures that make the environment perceptually accessible and enough to lose as little phenomenologically relevant information as possible. The cortical strategy could be adequate if it would consist of exploiting the ordering in the way natural objects appear that is contingent upon the perceptual access of an organism though it does not happen to be accidental (Witte 1958 gives an early mathematical treatment of the notion of perceptual orderings that fits experimental design). Accordingly, neuronal output correlations, which do not prove to be reducible, are deemed to subserve the recovery of orderings through grouping rules, and high level neuronal dependencies are accounted for as bearing the functional role of matching non-linearities due to the phenomenological structures occurring in natural scenes. Dependencies of neuronal responses implemented in long-range connections are hypothesized to subserve the tiling of the joint space of stimulus parameters that mirror the capability of grouping rules to interact in competitive or cooperative ways to allow for such features as boundaries, surfaces and self-enclosed objects to appear.

Kovács & Julesz (1993) and Kovács (1996) provide evidence of how grouping rules and phenomenological conditions, such as those underlying figure-ground structure or the invariant perceptual axes of objects, occur at local and global levels and interact at different dimensions given their capability of combining with one another in meaningful connections. Furthermore, Kovács (1996) claims that conditions for boundary and surface properties and their connections that account for unitary perceptual objects to appear are consistent with sparse coding models of shape perception.

As Koffka (1935) and Köhler (1939) suggested, perceptual units are shaped by the search for an optimal balance between the internal constraints imposed by the form of perception systems and the adaptive approximation to the external relevant properties of the environment, as is the case with the surface and volume forces that bind up the shape of an oil droplet into the water.

Accordingly, the optimization followed by the perceptual system faced with the statistics of natural scenes can be taken to rest on phenomenological rules that bind what it is to appear as a meaningful unit in the environment to satisfy

a relevant subset of the indefinite many structures of natural scenes. Hence, optimality criteria for perceptual processing in terms of data signal processing operations (Barlow 1961, Field 1994) can be evaluated by the phenomenological relevance of natural image data.

The phenomenological specification of environmental objects' properties could satisfy as well the request of minimizing the assumptions for the search for statistical regularities in natural scenes, as is claimed in the SC models, since it would be reduced to the relevant dimensions that match the way objects appear.

On these theoretical and empirical grounds, a conjecture can be made that the Gestalt theory approach can be used to develop phenomenological models of how the perceptual system is supposed to extract meaning from the dependencies required by the structure properties displayed in natural scenes.

In the next paragraph, a proposal for a formal framework of phenomenological models is outlined that can be applied to natural scenes. The formal treatment is likely to make primitives and relations comparable to questions and findings in other fields of research, though at the expense of some loss of intuitiveness.

Particular attention will be paid to the conditions of the emergence of features that account for unitary perceptual objects to appear, and to phenomenological relations underlying the capability of features to combine with one another at various orders and dimensions in the perceptual scene, whose correlate could be identified with the integration of parameters in an abstract multidimensional stimulus space.

Therefore it can be claimed that theoretical and experimental Gestaltist research, whose principles and findings are translated in a formal framework, can supplement the abstract definitions of stimuli as multidimensional spaces. Albrecht, Geisler & Crane (2004) argue correctly that the use of abstract multidimensional spaces is needed to avoid misleading definitions of what a stimulus should look like. But that would apply only to the apparent intuitiveness of introspective data.

Indeed, the Gestalt theory of perception as a self-organizing process that tends optimally to approximate the structures of the environment by constraints over the perceptual system that obtain extremal minima or maxima values does not rein in giving its principles and findings a probabilistic formulation. Furthermore, the formal translation of some Gestaltist findings for a phenomenological modelling does not commit to top-down, bayesian ideal observer or bottom-up probabilistic analyses. It can be used to qualify the assumptions that are embedded in the hypotheses whence the optimization decision rules and priors are derived. And a criterion can be further developed by which what counts as a feature may not be merely inferred from data through priors that happen to rest upon *ad hoc* association rules or by-hand criteria of computational form.

## 6. Phenomenological Modelling

The structure in natural scene images is deemed to be meaningfully correlated to environmentally relevant and salient properties for the behaviour and cognition of perceiving agents. But the structure can narrow to spatio-temporal forms of data ordering without any reference to the environment being implied. Perceptual reference is introduced when an agent is specified who picks up the information that the structure conveys under the conditions of rules or clues that are derived from the way the physical and geometrical properties of light and surfaces are realized in optical structures (see Koenderink & van Doorn 2003 who make noteworthy reference to the notion of “intentionality”).

In fact, rules and clues can be variously construed in many different theoretical frameworks. But Gestalt psychology, that is an empirical theory that realizes a form of phenomenology, takes them to be a sort of non-overt know-how of the effectual ways to parse the environment in meaningful units as wholes made up of parts that hold one another with many perceptual relations of mutual integration or segregation, requests of completion, dependence or mutual exclusion.

Hence, it is assumed that manifold chunks of image structures mean qualitatively different amounts of information for perceiving agents, which gives rise to *phenomenological structures*. That is rules, clues and conditions that make distinct sections of the environment phenomenally accessed for perceiving agents at different scales but always within the level at which agents and the environment are comparable, since the various environmental perceptual properties derive their meanings from driving purposive behaviour (Koffka 1935; Metzger 1972; Gibson 1979).

Grouping rules and phenomenological conditions are assumed to allow the environment perceptually to appear in a determinate phenomenological variety.

By phenomenological conditions are meant both limited conditions for a determinate phenomenon such as amodal completion to occur and more general phenomenological principles that apply to natural objects such as those stated by Musatti (1926). According to Musatti, the cognitive distinction between a perceptual aggregate of elements and a whole unified natural object is explained by such principles as the closure of possible transformations of colour or shape data, the data consistency in their simultaneous or sequential presentation, the continuity of transformation and change in apparent properties of objects in the different circumstances where they happen to be perceived, and the data belongingness to a whole unit, along with independence and completeness clauses.

Gestalt theory usually describes phenomenological structures with features, relations and composition rules, all of which can be modelled in terms of part-whole mereology that is integrated with such intuitive topological properties as



boundaries, connections that are needed to tell the aggregate made up of variously related parts from the “all of a piece” whole as a perceptual unit made up by the connection of only and all its parts (Varzi 1994a, 1994b and 1996 give an early formal treatment of such mereotopological issues). Part-whole properties of the standard Gestalt theoretical approach and experimental design can be described as ruled by models of extensional mereology axioms (Simons 1987) with a refinement that assumes as a primitive notion the definition of “proper part” along with the topological conditions for something to appear as a natural object (Smith 1996). In particular, the Gestalt conditions at which the qualification of a unit as a whole holds are theoretically analysed by Rausch (1937) and (1964), Rescher & Oppenheim (1959).

In Cognitive Sciences, the meaning of such terms as “model” and “modelling” may indeed differ. But they appear to be used mostly with reference to a conceptual theory and a computational description of a cognitive function, an interpolation of sets of data of various sorts, a functional architecture underlying a cognitive function that is implemented in biological and artificial systems. In fact, Suppes (1969) suggests that in empirical sciences are found examples of modelling that comply with a concept of models that is the same as the mathematical concept of model-theory. Therefore, the assumption can be made that perception research can profit from such convergence in specified cases, and that phenomenological structures can be regarded as a qualified instance of these cases.

Spillmann (2009) reports some phenomenological structures that can model, that is conceptualize, as many neuroscientific data of specific visual cortical neuron responses. It is noteworthy that the structures derive from experimental evidence provided by either experimental phenomenology, that is a research that is bound to experimental design where all the variables are observable (Bozzi 1989, 1999; Kanizsa 1979; Sinico 2003), or Gestalt psychology that admits instead a neurophysiological interpretation.

Spillmann lists the following structures. Grouping rules, which allow for unifying parts of consistent objects even though there are intersections of disjointed parts of distinct objects, model V1, V2, V3 neurons tuned to co-axiality, collinearity and smooth continuous motion trajectories of visual elements. Conditions for amodal completion with occlusions that meaningfully correlate as early as with V1 neurons.

The relation of border ownership, which qualifies a boundary as a margin, and of surface belongingness that provide models for V1 and V2 neurons that respond selectively more for the margin of a figure than for the ground, as well as for V2 neurons whose responses correlate with relative depth information due to figure-ground segregation. The conditions for abutting gratings, bi-dimensional or tri-dimensional Kanizsa surfaces without luminance gradients, which are founded



on well-defined phenomenological properties of inducers, that model V1 and V2 neurons. Apparent transparency segmentations of coplanar monochromatic partially overlapping shapes, whose topological, figural and colour conditions provide models for V2 neurons that assign properly borders to the not occluded surfaces that emerge from transparency.

But phenomenological structures, be they grouping rules, clues or conditions, can provide models in the mathematical sense, with the proviso that the usual set-theoretic foundations be replaced with the formal mereotopological definitions that fit more adequately such phenomenological primitives as parts and their relations (i.e. overlap), or to relations such as closure, connectedness, and of dependence and foundation that account for the capability of unilateral or mutual integration among various dimensions of perceptual properties and units.

In that sense, given a data set  $S$  from natural images, a phenomenological structure  $I$  is a model of  $S$  if it recovers the information that makes  $S$  accessible to the observer for a specified mereotopological property or a condition systematically coupled with a well-defined appearance of a natural object or of one or many of its perceptual properties.

The model  $I$  consists of features (such as boundary or surface) which match what information about parts and properties of objects is recovered by  $S$ , and of mereotopological relations, which specify how parts and properties are connected into an integrated “all of a piece” whole. Features and relations make up the  $I$  structure. Of course, the class and numerosity of what is a feature or a relevant relation vary a lot as a function of the determinate structure involved. Natural scene data are the domain or the universe of  $I$ s. The domain is many-sorted if it implies features or relations with different orders and spanning different dimensions. Each  $S$  gives a partition of phenomenological structures into those which are  $I$  models of it and those that are not. All the phenomenological structures that are  $I$ s of  $S$  can be termed as  $\text{Mod}(S)$ , that is the  $n$ -tuples structures that recover the perceptual information conveyed by  $S$ . That can be generalized as classes of natural scene data  $N$ .

Without going into further formal details, one can only say that the collection of features and relations of  $\text{Mod}(N)$  gives the phenomenological primitives for natural scenes in  $N$  that represent the information that is embedded in their image structures. Hence a phenomenological structure is a model that applies to a domain endowed with a specifiable relation, whose primitives are fixed by some relevant parameters (i.e. the grouping rules). If it is empirically shown that some phenomenological structures, which are  $\text{Mod}(N)$ , are also  $I$  models of other structures  $\text{Mod}(R)$ , whose data domain and primitives are different as is the case with neuroscience findings, then  $\text{Mod}(R)$  primitives are definable in terms of well-defined and tested phenomenological structures by some linking

propositions with determinate parameters. (Obviously other forms or relations among  $\text{Mod}(N)$  and  $\text{Mod}(R)$  obtain that are interesting for research).

In this framework, features of phenomenological structure match quantities of those subsets of relevant natural structures since they are regularities for the perceptual system. Regularities are invariant properties and conditions such that they preserve information as to the composition of units and their relations in image scenes through transformations due to perceptual system constraints and variations induced by the contingent perceptual circumstances, for instance translations, rotations and oclusions due to observer and object position and motion. A feature that matches a determinate quantity of visual structure is a  $n$ -placed variable whose values are realized in determinate series or sequences of patches in natural images. The value course is a magnitude, that is there is an ordering mereological and topological relation induced on a definite range of values. Each phenomenological magnitude can be linked to primitives of other models, as in the case of Neuroscience findings, or to abstract  $n$ -dimensional space stimulus maps.

Then, it can be argued that for determinate cases of phenomenological and Gestalt theory findings that are proved to be right, models are not deemed adequate to natural scenes by mere stipulation or a priori assumptions, since their primitives recover how features are distributed over and relations connect determinate natural image parts.

Therefore, primitives in ecologically valid visual domains are definable by a decomposition induced by a relevant transformation coupled with an equivalence relation. Each  $n$ -placed feature variable is identified with the invariant of at least  $2n$  terms equivalence relation. Every feature belongs to a definite order which is at least  $n-1$  smaller with reference to the visual structure of order  $n$  of which it is perceived as a part. That order determination is perceptually defined through the results of relevant transformations which hold dependencies constant, along parameters designated by grouping factors. The order is then derived by closure and connection properties of that which qualifies as a visual unit.

For instance, let the 2-placed variable  $B$  be the visual feature *boundary-for*, the  $n$ -sequence of data points in  $I(x,y)$  be the corresponding  $n$ -domain of locations  $(a_0, \dots, a_n)$ , and the  $B$  quantity distribution in  $(a_0, \dots, a_n)$  be the course of values of  $B(x,y)$  units in that domain.  $E(B)$  is an equivalence relation if for chains  $(b_0, \dots, b_{n-1}) \wedge (b_{n-1}, \dots, b_n) \wedge (b_n, \dots, b_{n+1})$  that are all sub-sequences in  $(a_0, \dots, a_n)$ , it assigns to all units in the sequences the same value satisfying  $B$  for a given transformation  $T$  that affects a relevant grouping parameter, that is for instance: closure, smooth continuity and convexity.

The located course of values in the visual field holds  $B$  constant if the visual bounding process is such that

$$\partial B(a_n) \sum_{i=0}^n (-1)^i (a_0, \dots, a_n)$$

where  $\partial B$  is the boundary visual extraction over the domain of locations, and  $(-1)^i$  is a component  $S$  of a dimension at least  $r$  with ordering  $r > p(a_n)$  such that for  $B = \Sigma \partial(a_n)$ ,  $(a_n) \in S$ ; while  $T$  is any transformation that maps values of  $a_n$  such that (a) any two of them in the subsequences  $b(b_0, \dots, b_n)$  are not equivalent to those occurring at other subsequences  $x(x_0, \dots, x_n)$  that do not lie within  $a_n$ ,

(b) for any discriminable visual area  $s^\circ$  of the interior region  $S^\circ = S_i^{(r)} / \partial B_n^{(p)}$  there is no neighbourhood  $S^\circ(Ns)_i^{(r)}$  that with every  $X^\circ(Nx)_i^{(r)}$  of a region  $x$  for another  $X^\circ = X_i^{(r)} / \partial B_n^{(p)}$  gives a product that is  $\neq 0$ .

Notice that this boundary invariance condition is indeterminate as it is symmetrical or asymmetrical in nature. It states only that a boundary is a difference-part with respect to the interior of a region and allows for it to be shared by different closed regions. Hence for invariances related to the border or margin, that is the features that segregate mutually and unilaterally two units into a circumscribed closed figure and a separate ground (Rubin 1921), further qualifications are needed which require further topological notions to be explicitly introduced (Smith 1996).

With further specifications, it is possible to model the phenomenological differences that decompose natural scene patches in relevant and perceptually salient units, disambiguating what may seem a cluttered array of edges, line terminations or contour segments. In the model, a specified feature  $F$  for the  $(a_n)$  domain is defined as the  $n$ -tuple of the induced  $n$ -fold cartesian product that satisfy the course of values which holds  $F$  constant. Since recognition of functional parts in a unitary and connected whole comes with decomposition through transformations, the assignment to a common bounding unit  $B$  of the visual meanings of a contact point among surfaces, a boundary or a margin, an occlusion or inclusion region helps the perceptual system to capture the natural image structure. On the grounds of phenomenological models, it is possible to study how dependencies and interactions of neuronal responses and long-range connections allow for the visual system to match structures in natural images.

Even at this stage it can be seen that feature magnitudes are connected in various dimensions of different order in the stimulus space. A boundary is a 2-placed variable, since its occurrence requires an interior region to be delimited, of 1-order which belongs to a surface, that is a structure of 2-order. The same applies for the smallest visually discriminable segments, which are all 0th-order and belong to boundaries or borders (i.e. the visual parts nearest to the boundaries), and for surfaces that have to occur as lower order units belonging to higher order visual units as solids of which they are faces.

Theories and experimental results in phenomenology and Gestalt psychology individuated two families of momentous relations, which allow the nature of that belongingness relation to be determined among units of different orders: dependence and foundation.

In a wide sense, dependence is captured by the following functional relation (Grelling 1939/1988). A variable  $f$  depends on a class of variables  $\Phi$ , if when every  $\Phi$  takes the same value at the arguments  $x_1$  and  $x_2$ , then  $f$  must take *equal* value at  $x_1$  and  $x_2$ . Intuitively, this definition captures the idea of the variation of a quantity that brings to bear on the variation of another. It can be generalized as *equidependence*: if a variable  $g$  that belongs to a class  $\Phi$  of relevant factors for a phenomenon takes the same values at arguments  $x$  and  $y$ , then  $f(x) = f(y)$ . No matter why and how much factors change that are external to  $\Phi$ , if functions belonging to  $\Phi$  are made constant, then the values of  $f$  are kept constant too. Converse relations of *equivariability* cover different sorts of variability: from variability of different degrees, to co-difference or simultaneous variability, and *ceteris paribus* variability (Simons 1998 mentions all those cases and expounds some constraints to impose on equidependence in order to avoid trivial cases).

If a variable keeps values constant without being affected by variations in a family or a restricted class of features, then it is *functionally independent*.

When it is applied to elements bounded in a structure, dependence accounts for them to be perceived as constituents of a configuration, that is a whole whose composition satisfies a determinate relation. Rescher & Oppenheim (1959) state that if  $R(p_0, \dots, p_n)$  is an ordered sequence satisfying  $R$ , then  $\phi$  is a dependence relation if it obtains that the value of  $p_0$  in some feature  $\gamma$  is connected to the  $\Gamma$  values of  $p_{i < n}$  such that for every  $R$ -structure the  $\gamma$ -values of the first  $p_i$  are functions of the  $\Gamma$  values for all  $p_n$ .

In the case that configuration consists of just  $p$ , dependence is thought of as a law-like connection holding among a  $\gamma$ -value of its own and all other values of some qualified restricted feature class. For instance, such a relation can hold among the value of a series of units grouped as a single contour and the statistical distribution of some such parameters as collinearity, co-axiality and minimum trajectory deviation, whose variation constrains all the units to be counted in the contour feature class.

If  $\Gamma$  is reduced to a single  $\gamma$ , a dependence holds in that the value it takes for  $p_0$  is law-like connected to values of  $(p_1, \dots, p_i, \dots, p_n)$ . That is the case when values for all points of a surface without luminance gradients for margin contours are all the same as the others belonging to the same surface relative to the inducers for amodal completion to obtain (Kanizsa 1974). The same applies to well-known cases such as the Müller-Lyer figure, where  $\gamma$ -value is apparent length and  $p_n$  are adjacent angular sectors with respect to their position.

In a narrower sense, when supplemented by the foundation relation, dependence rules composition among parts of different orders within and across stimulus space. Foundation is a primitive relation assumed in phenomenological models, which is meant to express the inherent connection among structure elements based on their need for integration. If an instance  $f$  of a feature class  $\Phi$  of  $n$ -order at parts  $p_i$  for a set of arguments  $x$  takes determinate  $\phi$ -values in the range allowed for by instances  $g_i$  at  $p_j$  and if  $p_j$  are the smallest area where these  $\phi[f_i(x)g_i(x_n)]$  values are connected with instances  $\psi[(u_i(x)v_i(x_n))]$  of a feature class  $\Psi$  of *equal* or *different order*, then each instance of the first class, say a boundary, is founded in every instance of the second class, say a surface. In this sense, a boundary will be said to depend on a surface in a slightly new and meaningful way: boundary and surface must be integrated, since p-appearances as object components need features of different dimensions and order to be inherently connected.

Those feature properties and relations that subserve structures in phenomenological models can be taken to cover empirical facts that are results of independent experimental research.

Lamme (2004) summarizes different sorts of contextual modulation of V1 responses. First order effect is masking and pop-out configuration, which alters perceptual salience of such a single unit as a line segment shown in isolation or embedded in texture with conditions of similar vs. different, and spatially uniform vs. random elements. Higher order effects are: unidimensional segregation of elements aligned in an integrated contour; figure-ground segregation through visual ascription of surface belongingness to boundary segments; two-dimensional segmentation of segments at juxtaposed orientations in boundary and surface.

These findings can be mapped onto models that have phenomenological manifolds with dependence relation among different stimuli dimensions, and foundation that accounts for connection across different visual magnitude orders. It is worth noting that perceptual saliency at a determinate stimulus order of relevant factors for neuronal responses may not be preserved at higher orders. In figure-ground segregation, line segments at the centre of the figure elicit the same figure-modulated neuronal response as that triggered by elements at borders, even though they are flanked by segments at orthogonal orientation. Anyway, this modulation does not occur with segments located immediately outside the border. Therefore, neuronal responses seem to override the grouping factors ostensibly at work at local or at a lower order dimension of stimuli, while they reflect the higher level structure. Furthermore, Lamme (1995) and Zipser et al. (1996) reported a striking invariance of some cells to cues linked to figure-ground segregation and that responses are not additive since they are the same when cues are presented alone or in combination.

It can be suggested that those different response behaviours for various

stimuli dimensions and forms of stimuli combinations are accounted for by different ordered structures integrated by a foundation relation provided by phenomenological models.

Another useful treatment of principles underlying phenomenological structures is given by Lewin (1941) with his general framework for dependence relations that obtain through part-whole mereology and topological transformations.

To be sure, Lewin meant to model interdependence systems which may be considered as a special case of dependence structures when the value a feature takes at a part of a whole is constrained by the value distributions at all the other connected parts. Gestalt psychologists of the Berlin School contended it to be the main exemplification of phenomenal structures, which can be dubbed a determinational system when value distribution is a function of inner and outer boundary conditions that constrain systems to take extreme values as a stationary state (Köhler 1920; Koffka 1935; Grelling & Oppenheim 1937-8b/1988). But that framework may be taken to hold more generally.

Lewin describes part-whole decomposition and the involved dependence relation in terms of topological transformations. On the other hand, topological properties of structures are given a hierarchical ordering by mereological properties. It can be suggested that an appropriate perceptual construal of those of Lewin's principles of modelling allows one to build models to specify both stability and articulation of objects as they usually appear in ecologically valid viewing conditions.

Dependence is defined as that transformation change that does not preserve the equidependence as to a determinate feature class, say a boundary, of both units *As* and units *Bs*. In other words, *As* and *Bs* are dependent on one another if a change of state of *As* brings about a change of state of *Bs*.

Accordingly, independence is captured as a converse relation provided that it exceeds a determinate quantity  $\epsilon$ , that is an arbitrarily small quantity of variation under which the change induced by a determinate transformation of *As* does not affect mereological and topological properties of *Bs*, great though it may be.

In particular, Lewin's framework is meant to work out the conditions for models of interdependence. But it allows for variable degrees of connection that give rise to weak and strong dependence structures.

The main interest in Lewin's work lies in its capacity to model dependence relations within and among visual structures. Given the appropriate transformation whose application gives the distribution of values that assign some elements the function of parts, dependencies in visual structures may be mapped by curves that plot the relations between one part and all the others, or among ordered pairs or sequences of chosen parts. Accordingly, it can be used to quantify the nature, quantity and variability of determinate structure in a delimited region of a natural image patch.

Furthermore, Lewin's analysis can capture scale-invariant decompositions. At various scales, variations induced by a group of transformations may preserve equidependence among feature classes. But it proves equal to the task of specifying conditions at which there may be potentially different decompositions for defined values of dependence at multiple ordered quantities of change at orders  $k > s > m > w$ . The relativization of part-decomposition to threshold values at change scales may well account for the meaning conveyed by structure in natural images. Dependence value ranges of feature classes induced by a group of transformations at  $s$  may be less than the independence quantity  $\varepsilon$ . But that may not be the case at  $k$ . As a consequence, prevailing conditions in image structure at  $k$  are likely to lead the visual system to generalize or specify a decomposition whose properties are bound by the occurrence of new observable boundaries. The same applies for quantities above  $k$ . This property of the model can be exploited to study the visual system strategy for facing the problem of feature integration at different scales whose difficulties and ambiguities have been highlighted since Witkin (1983). In conclusion, phenomenological Gestalt models assume mereological and topological properties endowed with such relations as dependence and foundation in their various sorts as the real primitives of perceptual structures in ecologically valid conditions.

Models apply to experimental conditions, since they are instances of phenomenological properties that may be well specified, and translated in mathematical and probabilistic well-known forms.

On the one hand, mereology, dependence and foundation relations can be defined in matrix and vector representations, that which allows for neutral analyses of operations defined on them (Blecksmith & Null 1991). Given a universe  $U = (1, \dots, n)$  which admits elements of an image patch as members on which a relation  $R$  is defined, and a  $n \times n$  boolean matrix  $\mathbf{A}$  where  $[a_{ij}] \in \{0, 1\}$ , then let  $\mathbf{A}$  be called the incidence matrix of  $R$  whose entry values are 0 or 1 whether  $R$  respectively holds or not for each element  $i, j$ . Conditions on  $\mathbf{A}$  allow for defining properties which qualify  $R$  as one of the relevant mereological properties. Furthermore, if  $S$  is another relation on  $U$  and  $\mathbf{B}$  its incidence matrix, with implications including being a proper part, being disjointed or having common parts, then dependence and foundation can be expressed over matrices  $\mathbf{A}$  and  $\mathbf{B}$  for every pair of entries  $a_{ij}$  and  $b_{ij}$ , and the relative matrix product  $\mathbf{AB} = [c_{ij}]$  where

$$c_{ij} = \bigcup_{k=1}^n a_{ik} b_{kj} \quad \text{with} \quad \bigcup_{k=1}^n x_k = x_1 \vee \dots \vee x_{n-1} \vee x_n$$

With the introduction of such incidence vectors as the following one for the relation of a dependent feature whose perception requires it be part of a smallest phenomenal unit that connects it with another stimulus dimension



$$\mathbf{X} = \begin{bmatrix} x_1 \\ \dots \\ x_n \end{bmatrix} \quad x_i = \begin{cases} 1, \\ 0, \end{cases} \text{ if the mereological relevant relation } R_i \text{ holds or not}$$

conditions about correspondence between n-vectors and matrix extension allow for computing products between diagonal vectors and incidence matrices for dependence and foundation properties.

On the other hand, all mereological relations and topological properties of phenomenological models can be subject to the empirical inquiry of the probability of their degree of connection and distribution. For instance, equidependence of  $f(x_1, x_2)$  with a  $\varphi$ -class of features with respect to a probability density function holds when the  $P(x_2)dx$  for  $f(x_2)$  within  $dx$  in the environment of  $P(x_1)dx$  is a determinate  $P(x)dx$ , provided that every  $g_n(x_1, x_2) \in \varphi$  has the same values for  $x_1$  and  $x_2$  (Grelling 1939/1988).

The same case is made by Rescher & Oppenheim (1959) for measures of similarity of structures obtained by decomposition for relevant dependence and foundation relations through appropriate transformations. A class of features provides a magnitude endowed with a natural, that is non-contingent or merely stipulatory, topology. As already seen, an instance of feature is an n-tuple of values regarding position, dimension and order. Feature instances make up the components of visual structures according to such primitive relations as dependence and foundation as they are invariant through transformations. Given a group of transformations, a subgroup  $Z$  is associated with a homeomorphism among sequences of locations, and an isomorphism between mereological and topological properties. Hence a feature appears as an invariant for a determinate structure under this subgroup, that is its occurrences are equidependent in conditions that differ only by  $Z$ -transformation. The relevant subgroup refers to the parameters relevant for the well-known grouping factors. For example, segregation of dots in a lattice of positions due to proximity, such as in Wertheimer (1923), is an invariant feature subtended by an equivalence class obtained through  $Z$ -subgroups of contraction and dilation. Then the degrees of similarity of structures among assignments of feature values to domains of position can be measured statistically in many ways given the form of ordering and the nature of values admitted for it.

### Conclusions

It was argued that phenomenological models set a common ground upon which Gestalt theory, visual Neurosciences and computational modelling of natural scenes can compare data and interpretations for an integrated explanatory strategy for perception research to obtain. Phenomenological features and conditions that emerge as invariant given some relevant transformations in the environment



recover the perceptual information that natural images are supposed to convey to the observer.

Phenomenological models endowed with dependence and foundation relations, which are presumed to reflect law-like connections underlying phenomenological structures of the environment wherein the brain evolved, can be correlated to abstract multidimensional stimulus space in Neurosciences. Furthermore, they provide an ecological interpretation of the dependencies hypothesized for anatomical and functional connections. That meets the requirements to minimize a-priori assumptions, and to test neuronal units with ecologically valid settings that justified the introduction of the natural image scene analysis.

### Summary

Attempts to introduce Gestalt theory into the realm of visual neuroscience are discussed on both theoretical and experimental grounds. To define the framework in which these proposals can be defended, this paper outlines the characteristics of a standard model, which qualifies as a received view in the visual neurosciences, and of the research into natural images statistics. The objections to the standard model and the main questions of the natural images research are presented. On these grounds, this paper defends the view that Gestalt psychology and experimental phenomenology provide a contribution to the research into perception by the construction of phenomenological models for an ecologically meaningful interpretation of the empirical evidence and the hypothetical constructs of the natural image research within the visual neuroscience.

A formal framework for the phenomenological models is proposed, wherein Gestalt theoretical principles and empirical evidence are represented in terms of topological properties and relationships, which account for the order and structures that make the environment accessible to observers at a relevant behavioural level.

It is finally argued that these models allow us to evaluate the principles and the empirical evidence of various natures which are integrated from different fields into the research into perception, and in particular into visual neurosciences.

**Keywords:** Models, phenomenology, Gestalt psychology, natural images, visual Neurosciences.

### Zusammenfassung

Bestrebungen, die Gestalttheorie in das Forschungsfeld der visuellen Neurowissenschaften einzugliedern, werden sowohl auf theoretischer als auch experimenteller Ebene diskutiert. Um den Rahmen zu definieren, in dem diese Bestrebungen unterstützt werden können, werden in diesem Beitrag Merkmale eines Standardmodells skizziert, das als überlieferte Sicht der visuellen Neurowissenschaften und der Forschung zur Statistik natürlicher Bilder bezeichnet werden kann. Einwände gegen dieses Standardmodell sowie die Hauptfragen der Forschung zur Statistik natürlicher Bilder werden dargelegt. Auf dieser Grundlage verteidigt der Beitrag die Ansicht, dass Gestaltpsychologie und experimentelle Psychologie durch die Entwicklung phänomenologischer Modelle, die eine ökologisch bedeutsame Deutung empirischer Aussagen und theoretischer Konstrukte in der

Forschung zu natürlichen Bildern innerhalb der visuellen Neurowissenschaft ermöglichen, einen Beitrag zur Wahrnehmungsforschung liefern.

Für phänomenologische Modelle wird ein formeller Rahmen aufgestellt, innerhalb dessen gestalttheoretische Gesetzmäßigkeiten und empirische Nachweise in den Begriffen derjenigen topologischen Eigenschaften und Beziehungen dargestellt werden, die zu den Ordnungen und Strukturen beitragen, die eine Umgebung für die Betrachter auf einer relevanten Verhaltensebene zugänglich machen.

Schließlich wird dargelegt, dass diese Modelle die Evaluierung von Gesetzmäßigkeiten und empirischen Aussagen verschiedenster Art innerhalb unterschiedlicher Forschungsfelder der Wahrnehmungsforschung, speziell aber der visuellen Neurowissenschaften, ermöglichen.

**Schlüsselwörter:** Modelle, Phänomenologie, Gestaltpsychologie, natürliche Bilder, visuelle Neurowissenschaften.

The present paper was supported by the Miur-Università degli Studi di Palermo with the funding no. B71J09001300001 (2007-ATE-0136).

## References

- Albrecht, D.G. & Geisler, W.S. (1991): Motion selectivity and the contrast-response function of simple cells in the visual cortex. *Vision Neuroscience* 7, 531-546.
- Albrecht, D.G., Geisler, W.S. & Crane, A.M. (2004): Nonlinear Properties of Visual Cortex Neurons: Temporal Dynamics, Stimulus Selectivity, Neural Performance. In L.M. Chalupa & J.S. Werner (Eds.): *The Visual Neurosciences I*, 747-764. Cambridge (Mass.) - London: MIT Press.
- Albrecht, D.G. & Hamilton, D.B. (1982): Striate cortex of monkey and cat: contrast response function. *Journal of Neurophysiology* 48, 217-237.
- Atick, J.J. & Redlich, A.N. (1992): What does the retina know about natural scenes? *Neural Computation* 4, 196-210.
- Attneave, F. (1954): Some informational aspects of visual perception. *Psychological Review* 61, 183-193.
- Barlow, H.B. (1961): Possible Principles Underlying the Transformation of Sensory Images. In W.A. Rosenbluth (Ed.), *Sensory communication*, 217-234. Cambridge (Mass.): MIT Press.
- Barlow, H.B. (2001): Redundancy reduction revisited. *Network Computational Neuronal Systems* 12, 241-253.
- Bell, A.J. & Sejnowski T.J. (1997): The "independent components" of natural scenes are edge filters. *Vision Research* 37, 3327-3338.
- Blakemore, C. & Campbell, F.W. (1969): On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal images, *Journal of Physiology* 203, 237-260.
- Bleeksmith, R. & Null, G. (1991): Matrix representation of Husserl's part-whole-foundation theory. *Notre Dame Journal of Formal Logic* 32(1), 87-111.
- Bonds, A.B. (1989): Role of inhibition in the specification of orientation selectivity of cells in the cat striate cortex. *Vision Neuroscience* 2, 41-55.
- Bozzi P. (1989): *Fenomenologia sperimentale*. Bologna: Il Mulino.
- Bozzi P. (1999): Experimental Phenomenology: A Historical Profile. In L. Albertazzi (Ed.), *Shape of Forms*, 19-54. Dordrecht: Kluwer.
- Brunswik, E. (1956): *Perception and the Representative Design of Psychological Experiments*. Berkeley: University of California Press.
- Brunswik, E. & Kamiya, J. (1953): Ecological cue-validity of 'proximity' and of other Gestalt factors. *American Journal of Psychology* 66, 20-32.
- Campbell, F.W. & Robson, J.G. (1968): Application of Fourier analysis to the visibility of gratings. *Journal of Physiology* 197, 551-566.
- Carandini, M., Heeger, D.J. (1994): Summation and division by neurons in primate visual cortex. *Science* 264, 1333-1336.

- Carandini, M., Demb, J.B., Mante, V., Tolhurst, D.J., Dan, Y., Olshausen, B.A., Gallant, J.L., Rust, N.C. (2005): Do we know what the early visual system does? *The Journal of Neuroscience* 25(46), 10577-10597.
- Carandini, M., Heeger, D.J., Movshon, J.A. (1997): Linearity and normalization in simple cells of the macaque primary visual cortex. *The Journal of Neuroscience* 17(21), 6621-6644.
- Chandler, D.M. & Field, D.J. (2007): Estimates of the information content and dimensionality of natural scenes from proximity distribution. *Journal of the American Optical Society A* 24(4), 922-942.
- Daugman, J. (1980): Two-dimensional analysis of cortical receptive field profiles. *Vision Research* 20, 846-856.
- Daugman, J. (1988): Complete discrete 2D Gabor transforms by neural networks for image analysis and compression. *IEEE Transactions* 36(7), 1169-1179.
- De Angelis, G.C., Robson, J.G., Ozhawa, I., Freeman, R.D. (1992): Organization of suppression in receptive fields of neurons in cat visual cortex. *Journal of Neurophysiology* 68,144-163.
- De Angelis, G.C., Ozhawa, I., Freeman, R.D. (1993): Spatio-temporal organization of simple cell receptive fields in the cat's striate cortex. II. Linearity of temporal and spatial summation. *Journal of Neurophysiology* 69,1118-1135.
- De Angelis, G.C., Anzai, A. (2004): A Modern View of the Classical Receptive Field: Linear and Non Linear Spatiotemporal Processing by V1 Neurons. In L.M. Chalupa & J.S. Werner (Eds.): *The Visual Neurosciences I*, 704-719. Cambridge (Mass.) - London: MIT Press.
- De Valois, R.L., De Valois K.K. (1980): Spatial vision. *Annual Review of Psychology* 31, 309-341.
- De Valois, R.L., De Valois K.K. (1988): *Spatial Vision*. New York: Oxford University Press.
- De Valois, R.L., Albrecht, D.G. & Thorell, L.G. (1982): Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research* 22, 545-559.
- Ehrenstein, W.H. (2001): Perceptual Organization. In P.B. Baltes & N.J. Smelser (Eds.), *International Encyclopedia of Behavioral and Social Sciences* 16,11227-11231. Elsevier.
- Ehrenstein, W.H., Spillmann, L. & Sarris, V. (2003): Gestalt issues in modern neuroscience. *Axiomathes* 13, 433-458.
- Elder, J. H. & Krupnik, A. (2001): Contour grouping with strong prior models. In *Proceedings of the IEEE Conference on Computer Vision Pattern Recognition*, 414-421. Kauai (Hawaii): IEEE Computer Society Press.
- Elder, J. H. & Goldberg, R. M. (2002): Ecological statistics of Gestalt laws for the perceptual organization of contours. *Journal of Vision* 2, 324-353.
- Elder, J. H. & Zucker, S. W. (1993): The effect of contour closure on the rapid discrimination of two-dimensional shapes. *Vision Research* 33 (7), 981-991.
- Elder, J. H. & Zucker, S. W. (1996): Computing contour closure. *Proceedings 4th European Conference on Computer Vision*, 399-412. Cambridge (UK).
- Elder, J. H. & Zucker, S. W. (1998): Evidence for boundary-specific grouping. *Vision Research* 38, 143-152.
- Field, D.J. (1993): Scale-invariance and Self-similar 'Wavelet' Transforms: An Analysis of Natural Scenes and Mammal Visual System. In M. Farge, J. Hunt & J.C. Vassilicos (Eds.), *Wavelets, fractals and Fourier transforms. Developments and new applications*, 153-193. Oxford (UK): Oxford University Press.
- Field, D.J. (1994): What is the goal of sensory coding? *Neural Computation* 6, 559-601.
- Field, D.J. & Hayes, A. (2004): Contour Integration and the Lateral Connections of V1 Neurons. In L.M. Chalupa & J.S. Werner (Eds.): *The Visual Neurosciences II*, 1069-1079. Cambridge (Mass.) - London: MIT Press.
- Field, D.J., Hayes, A., Hess, R.F. (1993): Contour integration by the human visual system: evidence for a local "association field". *Vision Research* 33(2), 173-193.
- Field, D.J. & Tolhurst, D.J. (1986): The structure and symmetry of simple-cell receptive field profile in the cat's visual cortex. *Proc. R. Soc. London B* 228, 379-400.
- Fitzpatrick, D. (2000): Seeing beyond the receptive field in primary visual cortex. *Current Opinion in Neurobiology* 10, 438-443.
- Foldiak, P. (1995): Sparse Coding in the Primate Cortex. In M.A. Arbib (Ed.), *The Handbook of Brain Theory and Neural Networks*, 895-989. Cambridge (Mass.): MIT Press.
- Geisler, W. S. (2008): Visual perception and the statistical properties of natural scenes. *Annual Review of Psychology* 59, 167-192.
- Geisler, W.S. & Albrecht, D.G. (2000): Spatial Vision. In K. De Valois (Ed.), *Handbook of perception and cognition*, 79-128. New York: New York University Press.
- Geisler, W. S., Perry, J. S., Super, B. J., Gallogly, D. P. (2001): Edge co-occurrence in natural images predicts contour grouping performance. *Vision Research* 41, 711-724.
- Gibson, J.J (1979): *The Ecological Approach to Visual Perception*. Boston: Houghton Mufflin.

- Gilbert, C.D. (1992): Horizontal integration and cortical dynamics, *Neuron* 9, 1-13.
- Gilbert, C.D. (1993): Circuitry, architecture and functional dynamics of visual cortex. *Cerebral Cortex* 3, 373-386.
- Gilbert, C.D. & Wiesel, T.N. (1990): The influence of contextual stimuli on the orientation selectivity of cells in primary visual cortex of the cat. *Vision Research* 30, 1689-1701.
- Grelling, K. (1939/1988): A Logical Theory of Dependence. In B. Smith (Ed.), *Foundations of Gestalt theory*, 217-226. Wien: Philosophia Verlag.
- Grelling, K. & Oppenheim, P. (1937-8a/1988): The Concept of Gestalt in the Light of Modern Logic. In B. Smith (Ed.), *Foundations of Gestalt theory*, 199-205. Wien: Philosophia Verlag.
- Grelling, K. & Oppenheim, P. (1937-8b/1988): Logical Analysis of 'Gestalt' as a 'Functional Whole'. In B. Smith (Ed.), *Foundations of Gestalt theory*, 210-216. Wien: Philosophia Verlag.
- Hansen, B.C. & Hess, R.F. (2007): Structural sparseness and spatial phase alignment in natural scenes. *Journal of the American Optical Society A* 24, 1873-1885.
- Heeger, D.J. (1992): Normalization of cell responses in cat striate cortex. *Vision Neuroscience* 9, 181-197.
- Heeger, D.J., Simoncelli, E.P., Movshon, J.A. (1996): Computational models of cortex visual processings. *Proc. Nat. Acad. Sc. USA* 93, 623-627.
- Helmholtz, H. von (1867/1962): *Treatise on Physiological Optics III*. New York: Dover.
- Hubel, D.H. & Wiesel, T.N. (1959): Receptive fields of single neurones in the cat's striate cortex, *Journal of Physiology* 148, 574-591.
- Hubel, D.H. & Wiesel, T.N. (1962): Receptive fields, binocular interaction and functional architecture in cat visual cortex. *Journal of Physiology* 160, 106-154.
- Hubel, D.H. & Wiesel, T.N. (1977): Functional architecture of macaque monkey visual cortex. *Proc. R. Soc. London B* 198, 1-59.
- Hyvärinen, A. (2010): Statistical models of natural images and cortical visual representations. *Topics in Cognitive Science* 2, 251-264.
- Hyvärinen, A., Hurri, J., Hoyer, P.O. (2009): *Natural Image Statistics. A Probabilistic Approach to Early Computational Vision*. Dordrecht-Heidelberg-London-New York: Springer.
- Jones, J.P. & Palmer, L.A. (1987): An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *Journal of Neurophysiology* (58) 6, 1233-1258.
- Jung, R. & Spillmann, L. (1970): Receptive Field Estimation and Perceptual Integration in Human Vision. In F.A. Young & D.B. Lindsley (Eds.), *Early Experience and Visual Information Processing in Perceptual and Reading Disorders*, *Nat. Acad. Sc.*, 181-197.
- Kanizsa, G. (1974): Contours without gradients or cognitive contours? *Giornale Italiano di Psicologia* 1, 93-112.
- Kanizsa, G. (1979): *Organization in Vision. Essays on Gestalt Perception*. New York: Praeger.
- Kapadia, M.K., Westheimer, G., Gilbert, C.D. (2000): Spatial distribution of contextual interactions in primary visual cortex and in visual perception. *Journal of Neurophysiology* 84, 2048-2062.
- Koenderink, J. J. & Van Doorn, A. J. (2003): Pictorial Space. In H. Hecht, R. Schwartz, M. Atherton (Eds.), *Looking into Pictures*, 239-299. Cambridge (USA)-London: The MIT Press.
- Köhler, W. (1920): *Die Physischen Gestalten in Ruhe und im Stationären Zustand*. Berlin: Braunschweig. In W. Ellis (Ed.) (1992): *A Source Book of Gestalt Psychology*, 17-54, London: Routledge.
- Köhler, W. (1939): *The Place of Value in a World of Facts*. New York: Liveright Publishing Co.
- Koffka, K. (1915): Grundlegung der Wahrnehmungspsychologie. *Zeitschrift für Psychologie* 73, 11-90. In W. Ellis (Ed.) (1992): *A Source Book of Gestalt Psychology*, 371-379, London: Routledge.
- Koffka, K. (1935): *Principles of Gestalt Psychology*. New York: Harcourt.
- Kovács, I. (1993): Gestalten of today: early processing of visual contours and surfaces. *Behavioural Brain Research* 82, 1-11.
- Kovács, I. & Julesz, B. (1993): A closed curve is much more than an incomplete one: effect of closure in figure-ground segmentation. *Proc. Nat. Acad. Sci. USA* 90, 7495-7497.
- Lamme, V.A.F. (1995): The neurophysiology of figure-ground segregation in primary visual cortex. *Journal of Neuroscience* 15, 1605-1615.
- Lamme, V.A.F. (2004): Beyond the classical receptive field: contextual modulation of V1 responses. In L.M. Chalupa & J.S. Werner (Eds.): *The Visual Neurosciences I*, 720-732. Cambridge (Mass.) - London: MIT Press.
- Lewin, K. (1941): Analysis of the concepts whole, differentiation and unity. *University of Iowa Studies in Child Welfare* 18 (1), 226-261.
- Maffei, L. & Fiorentini, A. (1973): The visual cortex as a spatial frequency analyzer. *Vision Research* 13, 1255-1267.

- Maffei, L. & Fiorentini, A. (1976): The unresponsive regions of visual cortical receptive fields, *Vision Research* 16, 1131-1139.
- Mahamud, S., Thornber, K. K., Williams, L. R. (1999): Segmentation of salient closed contours from real images. In *Proceedings of the 7th IEEE International Conference on Computer Vision*, 891-897. Los Alamos (CA): IEEE Computer Society Press.
- Marcelja, S. (1980): Mathematical description of the responses of simple cortical cells. *Journal of the American Optical Society* 70, 1297-1300.
- Marmarelis, P.Z. & Marmarelis, V.Z. (1978): *Analysis of Physiological Systems: The White Noise Approach*. New York: Plenum.
- Metzger, W. (1972): The phenomenal-perceptual field as a central steering mechanism. In Royce, J.R. & Rozeboom, W.W. (Eds.): *The Psychology of Knowing*, 241-265. New York/Paris/London: Gordon and Breach.
- Movshon, J.A., Thompson, I.D., Tolhurst, D.J. (1978a): Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *Journal of Physiology* 283, 53-77.
- Movshon, J.A., Thompson, I.D., Tolhurst, D.J. (1978b): Receptive field organization of complex cells in the cat's striate cortex. *Journal of Physiology* 283, 79-99.
- Musatti, C.L. (1926): Analisi del concetto di realtà empirica. In Musatti, C.L. (1964): *Condizioni dell'esperienza e fondazione della psicologia*, 15-175. Firenze: Giunti-Barbera.
- Olshausen, B.A. (1997): A functional model of V1 horizontal connectivity based on the statistical structure of natural images. *Soc. Neurosci. Abstract* 23, 2363.
- Olshausen, B.A. (2004): Principles of Image Representation in Visual Cortex. In L.M. Chalupa & J.S. Werner (Eds.): *The Visual Neurosciences I*, 1603-1613. Cambridge (Mass.)-London: MIT Press.
- Olshausen, B.A. & Field, D.J. (1996): Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* 381, 607-609.
- Olshausen, B.A. & Field, D.J. (1997): Sparse coding with an overcomplete basis set: a strategy employed by V1? *Vision Research* 37, 3311-3325.
- Olshausen, B.A. & Field, D.J. (2006): What is the Other 85 percent of V1 Doing? In J.L. van Hemmen & T.J. Sejnowski, *23 Problems in System Neuroscience*, 182-214. Oxford (UK): Oxford University Press.
- Rausch, E. (1937): Über Summativität und Nichtsummativität. *Psychologische Forschung* 21, 209-289.
- Rausch, E. (1964): Einzelgegenständlichkeit als phänomenale Eigenschaft. *Psychologische Forschung* 28, 33-45.
- Reid, R.C., Victor J.D. & Shapley R.M. (1997): The use of m-sequences in the analysis of visual neurons: linear receptive field properties. *Visual Neuroscience* 14, 1015-1027.
- Reid, R.C. & Usrey, W.M. (2004): Functional Connectivity in the Pathway from Retina to Striate Cortex. In L.M. Chalupa & J.S. Werner (Eds.): *The Visual Neurosciences I*, 673-679. Cambridge (Mass.) - London: MIT Press.
- Rescher, N. & Oppenheim, P. (1955): Logical analysis of Gestalt concepts. *The British Journal for the Philosophy of Science* 6(22), 89-106.
- Ringach, D.L. (2001): Spatial structure and symmetry of simple-cell receptive fields in macaque primary visual cortex. *Journal of Neurophysiology* 88, 455-563.
- Rubin, E. (1921): *Visuell Wahrgenommene Figuren*. København: Glydenalske Boghandel.
- Ruderman, D.L. (1994): The statistics of natural images. *Network: Computation in Neural System* 5, 517-548.
- Sarris V. (2006): *Relational Psychophysics in Humans and Animals. A Comparative-developmental Approach*. New York: Taylor & Francis Group.
- Schwartz, O. & Simoncelli, E.P. (2001): Natural signal statistics and sensory gain control. *Nature Neuroscience* 4(8), 819-825.
- Shapley, R.M. & Victor, J.D. (1978): The effect of contrast on the transfer properties of cat retinal ganglion cells. *Journal of Physiology* 285, 275-298.
- Sigman, M., Cecchi, G.A., Gilbert, G.D., Magnasco, M.O. (2001): On a common circle: natural scenes and Gestalt rules. *Proc. Nat. Acad. Sc.* 98(4), 1935-1940.
- Simons, P. (1987): *Parts. A Study in Ontology*. Oxford: Clarendon Press.
- Simons, P. (1998): Gestalt and Functional Dependence. In B. Smith (Ed.), *Foundations of Gestalt theory*, 158-190. Wien: Philosophia Verlag.
- Sinico, M. (2003): On the foundations of experimental phenomenology. *Gestalt Theory* 25, 111-120.
- Smith, B. (1996): Mereotopology: A theory of parts and boundaries. *Data Knowledge Engineering* 20, 287-303.
- Spillmann, L. (1997): Colour in a larger perspective: the rebirth of Gestalt psychology. *Perception* 26, 1341-1352.
- Spillmann, L. (1999): From elements to perception. Local and global processing in visual neurons. *Perception* 28, 1461-1494.

- Spillmann, L. (2009): Phenomenology and neurophysiological correlations: Two approaches to perception research. *Vision Research* 49, 1507-1521.
- Spillmann, L. & Ehrenstein, W.H. (1996): From Neuron to Gestalt. Mechanisms of Visual Perception. In R. Greger & U. Windhorst (Eds.), *Comprehensive Human Physiology I*, 861-893. Heidelberg: Springer.
- Spillmann, L. & Ehrenstein, W.H. (2004): Gestalt Factors in Visual Neurosciences. In L.M. Chalupa & J.S. Werner (Eds.): *The Visual Neurosciences II*, 1573-1589. Cambridge (Mass.) - London: MIT Press.
- Spillmann, L. & Werner, J.S (1996): Long range interactions in visual perception. *Trends in Neurosciences* 19, 428-434.
- Suppes, P. (1969): *Studies in the Methodology and Foundations of Science*. Dordrecht: Reidel.
- Van Hateren, J.H. & Ruderman, D.L. (1998): Independent component analysis of natural image sequences yields spatiotemporal filters similar to simple cells in primary visual cortex. *Proc. R. Soc. London Biol. Sci.* 265, 2315-2320.
- Van Hateren, J.H. & Van Der Schaaf, D.L. (1998): Independent component filters of natural images compared with simple cells in primary visual cortex. *Proc. R. Soc. London Biol. Sci.* 265, 359-366.
- Varzi, A. (1994a): On the Boundary Between Mereology and Topology. In R. Casati, B. Smith, G. White (Eds.), *Philosophy and the Cognitive Sciences*, 419-38. Vienna: Hölder-Pichler-Tempsky.
- Varzi, A. (1994b): Things All of a Piece. In M. Somalvico (Eds.), *Atti del IV Convegno dell'Associazione Italiana per l'Intelligenza Artificiale*, 55-58. Parma: Università degli Studi di Parma.
- Varzi, A. (1996): Parts, Wholes, and Part-Whole Relations: The Prospects of Mereotopology. *Data & Knowledge Engineering* 20, 259-86.
- Vinje, W.E. & Gallant J.L. (2000): Sparse coding and decorrelation in primary visual cortex during natural vision. *Science* 287, 1273-1276.
- Vinje, W.E. & Gallant J.L. (2002): Natural stimulation of the non-classical receptive field increases information transmission efficiency in V1. *Journal of Neuroscience* 22, 2904-2915.
- Wertheimer, M. (1923): Untersuchungen zur Lehre von der Gestalt. II. *Psychologische Forschung* 4(1), 301-350.
- Wilson, H.R. (1999): Non-Fourier Cortical Process in Texture, Form, and Motion Perception. In P.S. Ulinski, E.G. Johns, A. Peters (Eds.): *Cerebral Cortex 13. Models of Cortical Circuits*, 445-478. New York: Kluwer.
- Wilson, H.R. & Wilkinson, F. (2004): Spatial Channels in Vision and Spatial Pooling. In L.M. Chalupa & J.S. Werner (Eds.): *The Visual Neurosciences II*, 1060-1069. Cambridge (Mass.) - London: MIT Press.
- Wilson, S., Geisler, W. S., Perry J. S. (2009): Contour statistics in natural images: Grouping across occlusions. *Visual Neuroscience* 26, 109-121.
- Witkin, A.P. (1983): Scale-space filtering. *Proceedings of the Eighth International Joint Conference on Artificial Intelligence* 2, 1019-1022.
- Witte, W. (1958): Einführung in die Mathematische Behandlung Psychologischer Probleme. In F. Dorsch, *Psychologisches Wörterbuch* 4, 390-454.
- Zipsper, K. & Lamme, V.A.F. & Schiller, P.H. (1996): Contextual modulation in primary visual cortex. *Journal of Neuroscience* 16, 7376-7389.
- Zhangloul, K.A., Boahen, K., Demb, J.B. (2005): Contrast adaptation in subthreshold and spiking responses of mammalian Y-type retinal ganglion cells. *Journal of Neuroscience* 25, 860-868.

**Carmelo Cali**, b. 1974, is researcher at the Università degli Studi di Palermo (Italy), where he teaches Psychology of Arts for the Faculty of Humanities and the Faculty of Architecture. His research interests are in Gestalt Psychology, Phenomenology, Psychology of Perception, Cognitive Sciences, A.I. and Robotics, Visual and Auditory Neurosciences. His work is mainly addressed to the theories of picture and music perception, the historical background of the Gestalt Psychology and Phenomenology, and the perceptual basis of social-cognitive skills in the human-humanoid interaction.

**Address:** Dipartimento di Scienze Umanistiche, Università degli Studi di Palermo, viale delle Scienze 12, 90128 Palermo, Italy.

E-mail: carmelo.cali@unipa.it