Simple Gestalt Algebra

Eckart Michaelsen[1] and Vera V. Yashina[2]

[1]FhG-IOSB, Gutleuthausstrasse, 1, 76275 Ettlingen, Germany
[2]Dorodnicyn Centre RAS, 40, Vavilov str., Moscow, Russia

SCIENCE AND TECHNOLOGY PUBLICATIONS

Abstract. The laws of Gestalt perception rule how parts are assembled into a perceived aggregate. This contribution defines them in an algebraic setting. Operations are defined for mirror symmetry and repetition in rows respectively. Deviations from the ideal case are handled using positive and differentiable assessment functions achieving maximal value for the ideal case and approaching zero if the parts mutually violate the Gestalt laws. Practically, these definitions and calculations can be used in two ways: 1. Images with Gestalts can be rendered by using random decisions with the assessment functions as densities; 2. given an image (in which Gestalts are supposed) Gestalt-terms are constructed successively, and the ones with high assessment values are accepted as plausible, and thus recognized.

1 Introduction

The Gestalt Algebra is meant to capture the laws of perception – as unveiled e.g. by Wertheimer [10] – in a formal way. Thus the algebraic nature of perceived gestalts can be soundly described and in the end coded on machines. A gestalt is always an aggregate composed of its parts. It means more, but is completely determined by its parts. With universal algebra there is a sound apparatus for capturing these ideas with more rigors.

1.1 Related Work

The notion of a maximum-meaningful gestalt is due to A. Desolneux [2] referring to statistical models for background clutter and foreground gestalts much similar to the ones discussed here. She actually refers again to D. Love, D. Marr and even H. von Helmholtz. Above we already mentioned the pioneering work of Wertheimer.

The universal algebra (for which our definitions below are a special case) has been thoroughly investigated by A. I. Malcev [4]. Here operations are allowed with more than 0, 1, or 2 aryties. Such algebras are called homogenous, as long as they work only on one common set. Bikhoff & Lipson [1] generalize to heterogeneous algebras working on several sets (called phylae) in order to include things like modules, vector-spaces, etc. An important specialization of that for image generation, understand-

Michaelsen E. and V. Yashina V..
Simple Gestalt Algebra.
DOI: 10.5220/0004393300380047
In *Proceedings of the 4th International Workshop on Image Mining. Theory and Applications* (IMTA-4-2013), pages 38-47
ISBN: 978-989-8565-50-1
Copyright © 2013 SCITEPRESS (Science and Technology Publications, Lda.)

ing, and mining is the Ritter image algebra [8, 9]. A very sophisticated pattern algebra with generators and connectors has been defined by Grenander [3]. Own previous work – however rather informal, lacking proven results – has been published in [6, 7].

2 Definition of the Simple Gestalt Algebra

First the Gestalt space will be defined. Then two operations (|and ∑) will be given, constructing new elements from given ones. While | is a binary operation, ∑ is of unspecified arity. Closure will be proven, i.e. that all operations lead to well defined new elements of the Gestalt space. Thus the Gestalt algebra is a special form of a universal algebra in the Malcev sense. But we neither give neutral elements nor the projections, as is usually done in Algebra. Also inverse elements are not possible.

2.1 Gestalt Space

The product set

$$G = \mathbb{R}^2 \times \left(\mathbb{R} \bmod \pi\right) \times (0,\infty) \times [0,1] \tag{1}$$

will be regarded as Gestalt space (of the plane). For each $g \epsilon G$ the first component is called its **position** $po(g) \epsilon R^2$; the second component is called its **orientation** $or(g) \epsilon R \bmod \pi$; the third component is called its **scale** $sc(g)>0$; the fourth component is called its **assessment** $0 \leq as(g) \leq 1$. The Gestalt space obviously is a smooth manifold with boundaries. Figure 1 shows some random instances of the gestalt space. They are depicted in a grey-tone according to their assessment with white meaning assessment zero and black meaning assessment one. Uniform random attributes where chosen with position from $[0,10] \times [0,10]$, scale from $(0,3]$, and assessment from $[0,1]$ respectively.

2.2 Forming a Mirror Gestalt from two Parts

Definition 1: (Mirror operation)**.** The binary operation | is defined as

$$g \mid h = \left(\frac{1}{2}\left(po(g)+po(h)\right), ori\left(po(g)-po(h)\right), \mid po(g)-po(h) \mid + \sqrt{sc(g)sc(h)}, a_{|}\left(g,h\right)\right). \tag{2}$$

The *position* of the new Gestalt results from averaging the positions of its parts. The orientation of the new Gestalt is a function *ori* of a 2D vector *v.* It is obtained by $ori(v)=\arctan(v_y/v_x)$ for $v_x \neq 0$ and set to $\pi/2$ if $v_x=0$. The new scale is the sum of the Euclidean distance between the positions of the first and second part plus the geometric mean of the two scales. The assessment $a_|$ is a geometric mean of a product of four partial assessments:

SCITEPRESS
SCIENCE AND TECHNOLOGY PUBLICATIONS

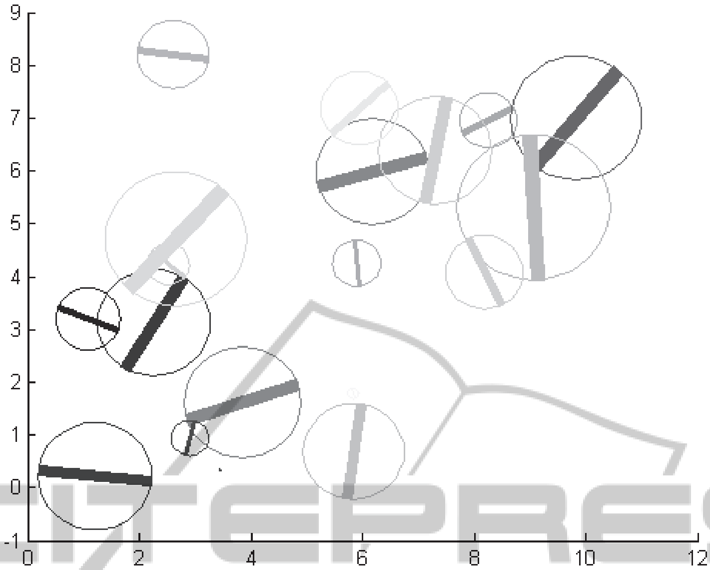

Fig. 1. A set of 20 randomly chosen gestalts.

SCITEPRESS
SCIENCE AND TECHNOLOGY PUBLICATIONS

$$a_{|}\left(g,h\right)=\left(a_{|,p}\cdot a_{|,o}\cdot a_{|,s}\cdot a_{|,a}\right)^{\frac{1}{4}}. \tag{3}$$

Before we give the definitions of these partial assessment functions let us consider the properties of two important help functions namely the **perceptual attention function** *exp(2-x-1/x)* on $x>0$ and the **angular neighbor function** $\alpha=½+½\cos(2x)$ on $0\leq x<\pi$ as they are displayed in Figure 2.

$$a_{|,p}=e^{2-\frac{|po(g)-po(h)|}{\sqrt{sc(g)sc(h)}}-\frac{\sqrt{sc(g)sc(h)}}{|po(g)-po(h)|}}, \tag{4}$$

has thus the form of a perceptual attention, and if a denominator in (4) should be zero we may set continuously $a_{|,p}=0$.

$$a_{|,o}=\alpha\left(or(g)+or(h)-2or(g\,|\,h)\right), \text{ where } \alpha \text{ is the angular neighborhood.} \tag{5}$$

$$a_{|,s}=e^{2-sc(h)/sc(g)-sc(g)/sc(h)} \text{ (also a perceptual attention),} \tag{6}$$

and another geometric mean

$$a_{|,a}=\sqrt{as(g)as(h)}\,. \tag{7}$$

Figure 3 shows examples of the operation at work. Actually, the small black part gestalts were generated using the method sketched in section 3.1.: 1) pick a random common orientation for the parts uniformly from $[0, \pi)$; 2) disturb the position, scale, and orientation attributes according to the ϵ value chosen. With rising ϵ the new gestalt (depicted in grey) will get a lower assessment (depicted as brighter grey).

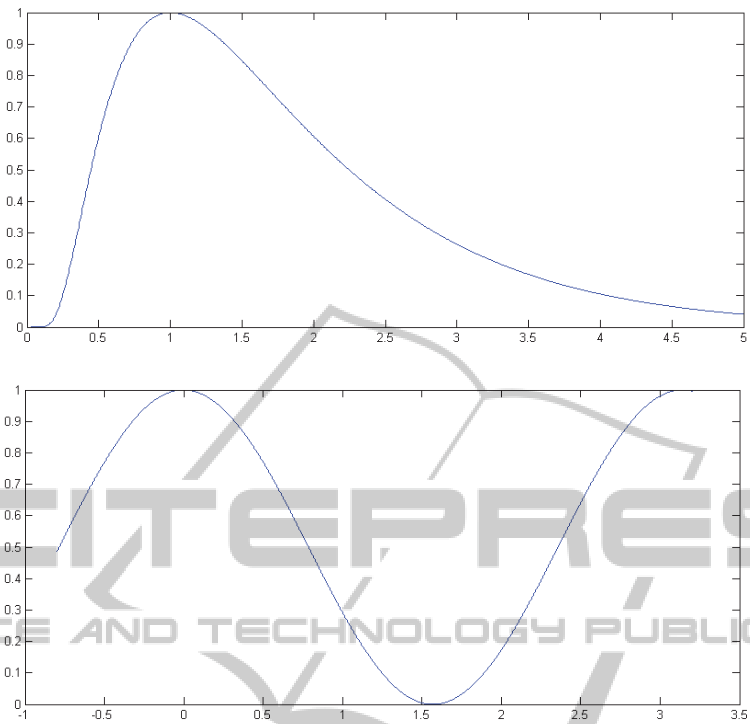

Fig. 2. Auxiliary functions for gestalt assessment calculation: Upper, perceptual attention;lower, angular neighborhood *α.*

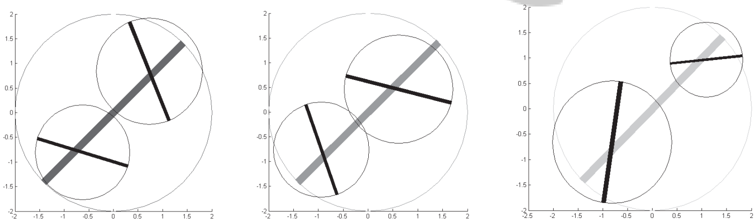

Fig. 3. Three terms of the mirror operation with left-to-right declining assessment ($\epsilon=0.05,0.1,0.2$).

Theorem 1. (Closure) $g|h \in G$ for any $g, h \in G$. Proof: 1) Position attribute: The average position of two points of $\boldsymbol{R^2}$ is in $\boldsymbol{R^2}$. 2) Orientation attribute: Arctangent yields always an orientation value modulo $\boldsymbol{\pi}$. Moreover, for $\boldsymbol{v_x} \rightarrow 0$ we have $ori \rightarrow \pm\ \pi/2$, so the function *ori* is smooth except for $v_x = v_y = 0$. This occurs e.g. for $g=h$; still also for these cases the orientation value is defined. 3) Scale attribute: Both, the absolute value, as well as the root, are positive, so is the sum of an absolute value and a root. 4) Assessment attribute: We shall prove that all four functions $a_{|,p}, a_{|,o}, a_{|,s}, a_{|,a}, \in [0,1]$ then the geometric mean (3) will also be there: $a_{|,p}$ and $a_{|,s}$ are functions of the form $exp(2\text{-}x\text{-}1/x)$; it is easily verified that this positive function takes its maximum value 1 for $x=1$; for $\boldsymbol{x} \rightarrow 0,\infty$ this function approaches zero (and thus is smooth everywhere,

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

see also Fig. 2); for $a_{|,o}$ consider (5) where some operations inside the orientation attribute are performed, which is a group with respect to addition; the function α is bounded between zero and one; $a_{|,a}$ is again a geometric mean of values from $[0,1]$.□

Lemma 1. (Commutativity) $g|h=h|g$ for any g, h. Proof: By checking symmetry of definitions and function *ori*. □ Note, that interchanging g and h can be interpreted as operation of the trivial two member group.

Lemma 2. (Self is worst) $as(g|g)=0$ for any g. Proof: By checking (4) we find $a_{|,p}=0$ for this case and thus a zero in the product (3).□

Remark 1. $a_{|,p}=1$ if $|p(g)-p(h)|$ equals the geometric mean of $s(g)$ and $s(h)$; else it will be smaller. If they are much further away of each other or much closer $a_{|,p}$ will approach zero. A deviation of some factor $t>1$ causes the same punishment as a corresponding deviation of factor $1/t$.

Remark 2. The orientation assessment function α gives $\alpha(0)=\alpha(\pi)=1$ and $\alpha(\pi/2)=0$; so according to (5) $a_{|,o}=1$ iff the orientations are mirror symmetric to each other with respect to $or(g|h)$. This is why we call this operation “mirror operation”. See also Figure 3 for examples of assessment better $1-\epsilon$.

Lemma 3. (Monotonicity) $(sc(g)\cdot sc(h))^{1/2}\leq sc(g|h)$ for any g, h. Proof: By checking (2) and because absolute values are positive.□

2.3 Summation into Row Gestalts

Definition 2: (Mirror operation)**.** The operation $\sum$ is of arity $n>1$. It is defined as

$$\sum g_1 ... g_n = \left(\frac{1}{n} \sum po(g_i), ori\left(\sum_{i=1}^{n-1} po(g_{i+1}) - po(g_i) \right), sc_{mid} + |\, po(g_n) - po(g_1)\,|, a_\Sigma\left(g_1,...,g_n\right) \right). \tag{8}$$

So the *position* of the new Gestalt results from averaging the positions of its parts. The orientation of the new Gestalt is obtained from summing up the difference vectors of the positions, where the function *ori* of a 2D vector v is again obtained by $ori(v)=arctan(v_y/v_x)$ provided that $v_x\neq 0$ and $ori=\pi/2$ else. The new scale is obtained from the Euclidean distance between the positions of the first and the last part plus the geometric mean of the scales $sc_{mid}=(sc(g_1)...sc(g_n))^{1/n}$. The assessment is again a geometric mean of a product of four partial assessments

$$a_\Sigma\left(g_1,...,g_n\right) = \left(a_{\Sigma,p} \cdot a_{\Sigma,o} \cdot a_{\Sigma,s} \cdot a_{\Sigma,a}\right)^{\frac{1}{n}}, \tag{9}$$

Where the positioning assessment is acquired as deviation from set-positions as

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

$$a_{\Sigma,p} = \left[\prod_{i=1}^{n} e^{-\frac{|po(g_i)-set_i|}{sc(g_i)}} \right]^{\frac{1}{n}} \text{ with } set_i = p(\Sigma g_1 ... g_n) + s(\Sigma g_1 ... g_n) \frac{i-(n-1)/2}{n-1} \begin{pmatrix} \cos(or(\Sigma g_1 ... g_n)) \\ \sin(or(\Sigma g_1 ... g_n)) \end{pmatrix} \quad (10)$$

and we make use again of assessing angular differences by function α of (4) setting

$$a_{\Sigma,o} = \left(\prod_{i=1}^{n} \alpha \left(o(g_i) - avo(g_1 ... g_n) \right) \right)^{\frac{1}{n}}, \quad (11)$$

Here the average orientation *avo* is obtained by summing up all orientations (as unit vectors) and using *ori* from (7). This can be problematic if the sum should equal zero.

$$a_{\Sigma,s} = e^{(2n - t_1 - ... - t_n - 1/t_1 - ... - 1/t_n)/n}, \text{ where } t_i = sc(g_i)/sc_{mid} \quad (12)$$

and sc_{mid} is again the geometric mean of the scales, and also

$$a_{\Sigma,a} = \left(a(g_1) \cdot ... \cdot a(g_n) \right)^{\frac{1}{n}}. \quad (13)$$

SCITEPRESS
SCIENCE AND TECHNOLOGY PUBLICATIONS

Figure 4 shows examples of the operation at work. Actually, the small black part gestalts were generated using the method sketched in section 3.1.: 1) Specify how many parts are to be built (in this case 5). From this follows the set-positions and the common scale of the parts; 2) pick a random common orientation for the parts uniformly from $[0, \pi)$; 3) disturb the position, scale, and orientation attributes according to the ϵ value chosen. With rising ϵ the new gestalt (depicted in grey) will get a lower assessment (depicted as brighter grey).

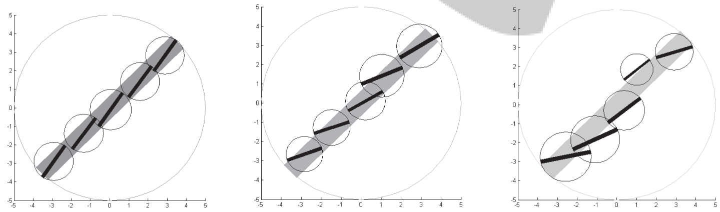

Fig. 4. Three terms of the row operation with left-to-right declining assessment ($\epsilon=0.05, 0.1, 0.2$).

Theorem 2. (Closure) $\sum g_1 ... g_n \in G$ for any $g_1 ... g_n \in G$. Proof: 1) Position attribute: The average position of n points of $\boldsymbol{R}^2$ is in $\boldsymbol{R}^2$. 2) Orientation attribute: Arctangent yields always an orientation value modulo $\boldsymbol{\pi}$ (see part 2 of proof of theorem 1). 3) Scale attribute: Both, the absolute value as well as the root are positive, so is the sum of an absolute value and a root. 4) Assessment attribute: We shall prove that all four functions $a_{\sum,p}$, $a_{\sum,o}$, $a_{\sum,s}$, $a_{\sum,a} \in [0,1]$ then the geometric mean (9) will also be there: $a_{\sum,p}$ is a product of functions of the form $exp(-x)$ where $0 \leq x$ and thus bounded by zero and one; $a_{\sum,o}$ is a geometric mean of values obtained by function α and thus bounded between zero and one; $a_{\sum,s}$ is obtained from (12) which may be decomposed into a product of n factors of form (4); $a_{\sum,a}$ is again a geometric mean of values from $[0,1]$.□

Lemma 4. (Generalized commutativity) $\sum g_n \dots g_1 = \sum g_1 \dots g_n$ for any $g_1, \dots g_n$. Proof: By checking symmetry of definitions and function *ori*. □ Note, the index mappings $\{i \to i, i \to n+1-i\}$ are a sub-group of the group of index permutations.

Lemma 5. (Self is worst) $as(\sum g \dots g)=0$ for any g. Proof: By checking (10) we find $a_{\sum,p}=0$ for this case and thus a zero in the product (3).□

Lemma 6. (Monotonicity) $(sc(g_1)\cdot \dots \cdot sc(g_n))^{1/n} \leq sc(\sum g \dots g)$ for any $g_1 \dots g_n$. Proof: By checking (8) and because absolute values are positive.□

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

3 Some Useful further Definitions and Results

For a small $\varepsilon>0$, we may define the relation $=_\varepsilon$ by $g=_\varepsilon h$ iff $|as(g)- as(h)|<\varepsilon$ and the other attributes of g and h are equal. Then we get

Lemma 7. For any small $\varepsilon>0$ exists a small angle δ such that for any g and h $g|h=_\varepsilon \sum gh$ if $|or(g)-or(g|h)+\pi/2|<\delta$ and $|or(h)-or(g|h)+\pi/2|<\delta$. Proof: Comparing (2) and (8) we find $po(g|h)=po(\sum gh)$, $or(g|h)=or(\sum gh)$, and $sc(g|h)=sc(\sum gh)$ respectively. The same holds for the assessment components: (4) equivalent to (10), (6) equivalent to (12), and (7) equivalent to (13). The only difference is in (5) versus (11). But for $|or(g)-or(g|h)+\pi/2|<\delta$ and $|or(h)-or(g|h)+\pi/2|<\delta$ we will have both $a_{\sum,p}>1-\varepsilon$ and $a_{|,p}>1-\varepsilon$.□

Recall here that for small angles δ the $cos(\delta)$ can be approximated as 1. So for very small $\varepsilon>0$, δ may be a considerable deviation. We will always have particular interest in such cases, where the same parts arranged differently in a term still yield the same (or ε-same) gestalt object:

Definition 3: (Gestalt equivalence**).** We define a gestalt-term recursively: 1) Each $g \epsilon G$ is a term. 2) For two terms s and t $s|t$ is a term as well as for n terms $t_1 \dots t_n$ $\sum t_1 \dots t_n$ is a term. Interchanging s and t as well as reordering $t_1 \dots t_n$ into $t_n \dots t_1$ does not change the value of the term in gestalt space (Lemmata 1 and 4). Thus there is an equivalence relation defined on the set of gestalt-terms. The corresponding equivalence classes are the main objects of our interest (gestalts).

Definition 4: (Search regions)**.** For $g\epsilon G$ and $\varepsilon>0$ the set $\{h\epsilon G;\ ass(g|h)>1-\ \varepsilon\}$ is called ε-mirror-search-region; For $g_i\epsilon G$, $0<i\leq n$, and $\varepsilon>0$ the union of sets

$$\bigcup_{j\neq i}\left\{g_j \in G; as\left(\sum g_1 ... g_n\right) > 1-\varepsilon\right\} \tag{14}$$

is called ε-i-n-row-search-region.

4 Gestalt Algebra at Work

We state that only gestalts with high assessments are meaningful. Thus the margin of the manifold, where $as(g)=1$, is of most interest. Close to this margin, i.e. where $as(g)=1-\epsilon$ with a small $\epsilon>0$, only small deviations from the ideal gestalt laws occur. Two possible utilizations are possible: Generative and reducing.

SCITEPRESS
SCIENCE AND TECHNOLOGY PUBLICATIONS

4.1 Generating Gestalts

One or more large Gestalts are set by a user filling the screen (or sheet). Then recursively each Gestalt is decomposed by: 1) Choosing randomly an operation from | and ∑ and (if it is not |) an arity n; 2) Choosing decompositions accordingly and at random such that the (normed) assessment functions are used as densities for drawing the new positions, orientations, and scales, respectively. To this end the integrals of the assessment functions over the whole admissible domain must be finite, so that they can be normed using the inverse of the integral as factor. Then they can be used as probability density functions for the random choice.

From Lemmas 3 and 6 it follows, that the Gestalts thus generated will become smaller and smaller. The generation may be terminated if they are smaller than a threshold (e.g. 5 Pixel). Then all these Gestalts can be drawn rendering an image. If the random numbers are not newly drawn for each branch of the term tree, but instead copied from one branch to the other, the image will exhibit interesting symmetries.

4.2 Reducing Gestalts

Primitive gestalts of small size are obtained from a digital image. The simplest way is using one of the commonly used edge detection methods. They are assessed e.g. according to the gradient magnitude. Position and orientation are also naturally given. The scale is fixed and obtains the same value for all primitives (such as two or three pixel). Another possibility is using the well-known SIFT method. It naturally sets the position, orientation, scale and assessment attributes for each extracted feature (i.e. primitive gestalt).

Then the gestalts are successively and at random combined into new gestalts. Again the assessment functions are used for drawing with preference. A suitable interpretation system for doing so has also been given by [5]. In the end terms of the

Gestalt Algebra are reduced from images. They may be visualized as reduction trees. Each such term has an assessment. For such procedure Definition 2 is of high interest, since equivalent terms may be reduced on different branches of the search. In order to avoid combinatorial explosion with the size of the terms, it is important to store the terms as representatives of the equivalence classes, and avoid multiple storage.

Drawing with preference means that the interpreter works on a set of existing gestalts. It selects better assessed ones with higher probability. Then it looks for suitable partners to form new gestalts using the three gestalt operations. "Suitable" here means it queries the set for partners such that the resulting gestalts have assessment $1\text{-}\epsilon$ with a small $\epsilon > 0$. **Search regions** in the gestalt space can be constructed by fixing ϵ and setting the derivative of the assessment function zero.

Figure 4 illustrates this: One gestalt g is displayed in the middle and possible partners h arranged around it such that the components of the assessment function are better than $1\text{-}\epsilon$. Possible partner gestalts are of about the same size, and in appropriate distance. The search region for the position is bounded by two concentric circles around g. Most important: Their orientation roughly fits the mirror symmetry constraint.

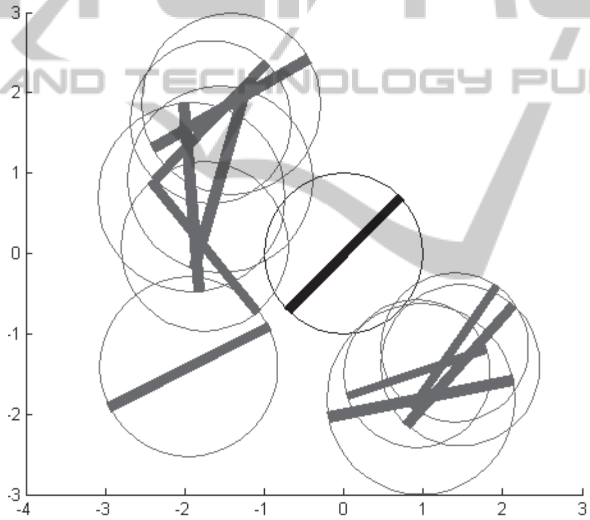

Fig. 5. A gestalt (black) and a set of 10 randomly chosen partner gestalts (grey) with position, scale and orientation attributes in the $1\text{-}\epsilon$ domain ($\epsilon=0.1$).

5 Discussion, Outlook, and Conclusions

We have presented an algebraic setting for the laws of gestalt perception. With such definitions, lemmas and theorems at hand the endeavor of machine coding such perceptive capabilities, as they are found in e.g. human subjects, will be facilitated. Much work lies still ahead:

5.1 Product into Rotational Mandalas

In [7] an operation $\prod$ was defined of arity $n>1$. For a good assessment the parts should be arranged in a regular polyhedron with n vertices. An iterative solution was given (equation (4)) for the attributes *po* and *sc*, fitting quite perfectly into the settings here. The position of the new Gestalt is initialized by averaging the positions of its parts. The radius of the polyhedron can be initialized from the mid distance of the parts to the initial position. A closure theorem would have to prove the convergence of the iteration.

We did not (yet) include this operation here, because there is a problem with the orientation attribute: The iteration outlined in [own citation] yields a *phase* defined in $[0,2\pi/n)$. This does not really fit the intentions of the second component of G in (1). The operation would be of interest, for instance because an interesting Lemma on generalized commutativity could be stated here stating $\prod g_1 \ldots g_n = \prod g_{\sigma(1)} \ldots g_{\sigma(n)}$ for any $\sigma \in D_n$ where D_n is the Dihedral group of n. So here we have non-trivial group action on the gestalts. We will investigate this further and modify the definitions such that this operation can also be included consistently.

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

References

1. Birkhoff G., Lipson J. D.: Heterogeneous Algebras. Journal of Combinatorial Theory, Vol.8, 1970. pp. 115-133.
2. Desolneux A., Moisan L., Morel J.-M.: From Gestalt Theory to Image Analysis. Springer, Berlin, 2008.
3. Grenander U.: General Pattern Theory. Oxford Science Publications, 1994.
4. Malcev A. I.: Algebraic Systems. Nauka, Moscow, 1970; Springer, Berlin, 1973.
5. Michaelsen E., Doktorski L., Luetjen K.: An Accumulating Interpreter for Cognitive Vision Production Systems. Pattern Recognition and Image Analysis, 2012, Vol. 22, No. 3, pp. 464–469.
6. Michaelsen E., Meidow J.: Basic Definitions and Operations for Gestalt Algebra. In: Gourevich I., Niemann H., Salvetti O. (eds.): IMTA 2009 – Image Mining Theory and Applications. Lisbon, 2009. pp. 53-26.
7. Michaelsen E.: On the Construction of Gestalt-Algebra Instances and a Measure for their Similarity. In: Gurevich I., Niemann H., Salvetti O. (eds.): IMTA 2010 – Image Mining Theory and Applications. Angers, May 2010. pp. 51-59.
8. Ritter G. X.: Image Algebra. Online at: http://www.cise.ufl.edu/~jnw/CVAIIA/ (accessed 30.12.2012), 1993.
9. Ritter G. X., Wilson J.N.: Handbook of Computer Vision Algorithms in Image Algebra, 2-d Edition. CRC Press Inc., 2001.
10. Wertheimer M.: Untersuchungen zur Lehre der Gestalt, II. Psychologische Forschung, Vol. 4 , 1923. pp. 301-350.