Representations of Posets to Generate Emerging Images

Agustín Moreno Cañadas, **Nelly Paola Palma Vanegas**
Department of Mathematics, National University of Colombia
Bogotá - Colombia
amorenoca@unal.edu.co, nppalmav@unal.edu.co

Keywords: Bot, Captcha, Emerging images, Gestalt, Representation of posets.

Abstract

In 2009, N. J. Mitra, H. Kuo Chu, T. Yee Lee, L. Wolf, H. Yeshurun and D. Cohen-Or proposed a synthesis technique to generate emerging images of 3D objects. Such images are gestalts with the property of to be detectable by humans but difficult to process by computer vision algorithms [1]. Therefore, emerging images can be used to design a blind test to tell humans and machines apart (CAPTCHA). In this paper, we use representations of a finite poset over a finite field in order to generate a large number of emerging images and different kind of gestalts.

1 Introduction

The representation theory of posets was developed in the 1970's by Nazarova, Roiter and their students in Kiev with the purpose of giving a solution to the second conjecture of Brauer-Thrall on classification of algebras [2–5].

According to Nazarova, Roiter and Gabriel. A *linear representation* U of a given poset $(\mathcal{P}, \leq)$ over an algebraically closed field k is a system of subspaces of a fixed vector space U_0 with the form:

$$U = (U_0, U_x \mid x \in \mathcal{P}), \tag{1}$$

where for each $x \in \mathcal{P}$, U_x is a subspace of U_0 such that $U_x \subseteq U_y$ provided $x \leq y$.

For a given poset $\mathcal{P}$, the main goal of the poset representation theory consists of giving a complete description of all indecomposable objects of the additive category rep $\mathcal{P}$ of representations of $\mathcal{P}$.

This paper is devoted to a novel application of the poset representation theory. Actually, we use representations of posets over a finite field to obtain a matrix model which generates a large number of emerging images. According to Mitra et al, *Emergence* is the phenomenon by which we perceive objects in an image not by recognizing the object parts, but as a whole, all at once. Although such phenomenon popularized by the Gestalt school has been well studied. The exact process of how we perceive such objects is not known up to date [1]. This fact allows to conclude that, it is extremely challenging, if not impossible, to automate the recognition process. For this reason, emergence can be used to make a good blind test, also known as CAPTCHA to distinguish between a human and an automated agent, commonly referred to as a bot [6, 7].

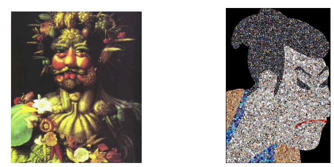

Figure 1: (left) Arcimboldo's painting, (right) Kabuki mosaic with 4200 tiles [8]

Mitra et al, hypothesized that emergence images are hard for automatic algorithms to segment, identify, and recognize taking into account the computational complexity of the human process required to detect objects in such images [1, 9]. In accordance with the Gestalt theory, the object *emerges* only when the relevant parts are exposed together giving the impression of the whole object. For example, Arcimboldos's paintings illustrate emergence by taking natural elements as vegetables and fruits, and juxtaposing them to suggest human and other shapes. *Digital image mosaics* are also built upon emergence. Mosaics are a form of art in which a large image is formed by a collection of small images called tiles. Various mosaics can be created for an image depending of the choice of tiles and the restriction in their placement. Tile mosaics, for example, are images made by cementing together uniformly colored polygonal tiles carefully positioned to emphasize edges in the composite picture (see Figure 1). We note that the investigation of digital mosaics has not been used to confuse bots, or prevent the use of computer vision techniques to potentially recover coherent object boundaries. In this paper, we describe the use of some suitable tiles to create emerging images [8].

The algorithm of Mitra et al was inspired by the well known image of the dalmatian dog by R.C. James, (see Figures 2 and 3), which is the best known demonstration of the emergence effect. According to them detecting the dog in this image is hard for humans, but definitely by far harder for a bot.

Figure 2: Dalmatian by R.C. James

In Figure 4 we show how a standard edge detector successfully extracts the feature curves of the original rendering whereas on the corresponding emerging image obtained by Mitra et al the results can be seen as noise [10, 11].

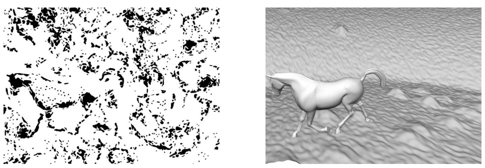

Figure 3: (left) Emerging image generated by the synthesized algorithm of Mitra et al.

According to L. Wolf et al, who have used emerging images to generate CAPTCHAS, the main limitation with emerging images seems to be the difficulty to create a large amount of recognizable models [12].

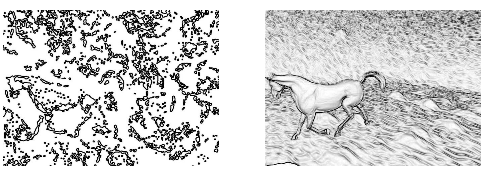

Figure 4: Edge detection of the emerging image shown in Figure 3.

In accordance with the Gestalt school, vicinity, similarity, continuity of direction, amodal completion, closure, constant width, tendency to convexity, symmetry, common motion, past experience, emergence, reification, multistability and invariance is the list of elementary grouping laws [9, 13, 14]. Such groups are identifiable with subsets of the retina. In image analysis we identify them with points of digital images. Whenever points (or previously formed groups) have one or several characteristics in common, they get grouped and form a new larger visual object, a *gestalt*. The famous Old-Woman/Young-Woman drawing (see Figure 5) is an example of multistability, such image consists in single set of lines with the help of them one can see two figures depending on how one's eye patterns or structures or interrelates those lines. The eye does not add anything to the image, but only organizes the lines presented to it, creating wholes out of disconnected experiences. The eye concocts configurations, and then it fits all the extraneous details into those configurations [15, 16]. In order to give advances to the problem proposed by L.

Figure 5: Old woman and young woman

Wolf et at, in this paper, we propose an algorithm which allows to generate a large number of *ambiguities* as defined for the law of multistability.

In [17], Cañadas and Vanegas introduced an extended visual cryptography scheme which uses a large number of nested halftone images and some biometric techniques to hide sets of gray-level images. Such scheme motivated us to define a generator of emerging images via admissible transformations of a matrix representation of a suitable poset. This new algorithm generates different kind of gestalts, therefore generalizes the model introduced by Mitra et al.

This paper is organized as follows; background and some related works are described in section §2, in §3 we present an algorithm to generate gestalts with the help of suitable representations of a poset, some experimental results are explained in §4, concluding remarks are described in §5.

2 Background and related works

2.1 Representations of posets

The representation theory of partially ordered sets or posets plays a prominent role in the representation theory of finite-dimensional algebras.

As we have said before, representations of posets were first introduced by L.A. Nazarova and A.V. Roiter. Such introduction was given by using matrix language in such a way that for a finite poset $(\mathcal{P}, \preceq)$ a representation of $\mathcal{P}$ over a field k is an arbitrary matrix

$$A = \begin{array}{|c|c|c|c|}\hline A_1 & A_2 & \dots & A_m \\ \hline \end{array} \tag{2}$$

with entries in k, partitioned horizontally into m (vertical) blocks (also called strips) [5]. Here the columns of a block A_i are formed by the coordinates (with respect to a chosen basis in U_0) of any minimal system of generators of U_i modulo its *radical* subspace $\overline{U}_i = \sum\limits_{j \preceq i} U_j$.

Let B another (matrix) representation of a poset $(\mathcal{P}, \preceq)$:

$$B = \begin{array}{|c|c|c|c|}\hline B_1 & B_2 & \dots & B_m \\ \hline \end{array} \tag{3}$$

A representation A is *isomorphic* to a representation B of a poset $(\mathcal{P}, \preceq)$ if A can be reduced to B by the following *admissible transformations*:

1. elementary transformations of rows of the whole matrix A;
2. elementary transformations of columns within each vertical strip;
3. additions of columns of a strip A_i to columns of a strip A_j if $i \preceq j$ in $\mathcal{P}$.

The *direct sum* of two representations A and B is the representation $A \oplus B$ which is equal to

$$A \oplus B = \begin{array}{|cc|cc|c|cc|} \hline A_1 & 0 & A_2 & 0 & \ddots & A_m & 0 \\ \hline 0 & B_1 & 0 & B_2 & \ddots & 0 & B_m \\ \hline \end{array} \tag{4}$$

Although, this paper is devoted to the applications of the poset representation theory in computer science, more specifically to the generation of emerging images, we must recall that the main problem of the poset representation theory is to obtain a complete description of the indecomposable objects (i.e. representations which cannot be written as a direct sum of two nonzero representations) in the category rep $\mathcal{P}$ for a given poset $\mathcal{P}$.

2.2 An algorithm to create emerging images

Inspired by the image of the dalmatian dog (see Figure 2), Mitra et al introduced an algorithm to create emerging images, such an algorithm renders 3D models by texturing them with large dots, which they called *splats*. Figure 6 shows some emerging images obtained with the algorithm of Mitra et al.

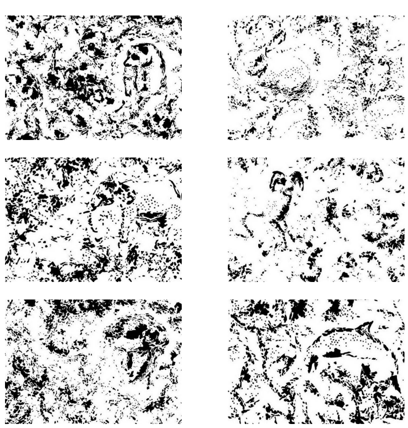

Figure 6: Bear, rabbit, elephant, goat, bust and dolphin as emerging images obtained via the synthesized algorithm of Mitra et al [1]

2.3 Nested images

Since the algorithm defined in this paper uses systems of nested images to generate emerging images. In this section, we provide a brief description of this type of schemes.

A nested image is a particular style of artwork, often but not exclusively based on silhouettes, in which an inner figure is nested in an outer figure, perhaps recursively [18]. The effect has a particularly interesting artistic effect if certain holes within an outer figure correspond to certain extremities of an inner figure, in a meaningful way. To produce an aesthetic result, the outer and inner figures must be designed carefully to ensure a good match. Figure 7 shows an example of nested images obtained via an algorithm defined by Tong et al. We note that nested images were used in [17] to hide halftone images.

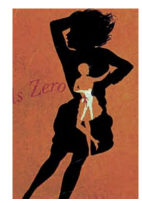

Figure 7: Example of nested images [18]

3 The process

In this section, we present an algorithm to generate gestalts. Such an algorithm can be carry out in three steps which we describe as follows:

1. Set a collection of nested emerging images $S = \{I_1, I_2, \ldots, I_k\}$ with $I_j = \sum\limits_{r_j=1}^{t_j} M_{r_j}$, M_{r_j} is an emerging image such that $M_{r_j} \subset M_{r_j+1}$ for all $1 \leq r_j \leq r_j - 1$ (i.e., S is a matrix representation over $\mathbb{Z}_2$ of a poset $\mathcal{P} = \sum\limits_{i=1}^{k} C_k$), C_k is a t_k-chain.
2. For each $j = 1 \ldots k$ write $M_{r_j} = \sum\limits_{U_{r_j} \in B_{r_j}} U_{r_j}$ where B_{r_j} is a tessellation for M_{r_j} (i.e., B_{r_j} is a set of generators of the vector space M_{r_j}) such that $B_{r_j} \subseteq B_{r_j+1}$.
3. Each image $M_{r_j} \in S$ can be obtained via admissible transformations of the matrix representation S.

4 Experiments and results

For the experiment, we have used 1200 images (extracted from well known galleries of fairy tales, famous circus acts and big cats) in order to fusion several systems of nested emerging images, such images have been scaled, translated and rotated to increase the number of images included in our scheme. Figure 10 shows examples of outputs obtained by our algorithm, using a Dell Precision M4400 PC (intel core 2 Quad Extreme Edition). It takes 5-30 minutes to obtain a complete system of nested images. In the proposed scheme, a system of nested images induces a representation U over $\mathbb{Z}_2$ of the following poset $\mathcal{P}$:

$$
\begin{array}{ccccc}
\circ_{a_1} & & \circ_{b_1} & & \circ_{c_1} \\
\uparrow & \nearrow & \uparrow & \nearrow & \uparrow \\
\circ_{a_2} & & \circ_{b_2} & & \circ_{c_2} \\
\uparrow & \nearrow & \uparrow & \nearrow & \uparrow \\
\circ_{a_3} & & \circ_{b_3} & & \circ_{c_3} \\
\uparrow & \nearrow & \uparrow & \nearrow & \uparrow \\
\circ_{a_4} & & \circ_{b_4} & & \circ_{c_4}
\end{array}
$$

Figure 8 shows the ground space U_0 in the representation U. In this case each image is obtained via admissible transformations over $\mathbb{Z}_2$ of the corresponding matrix representation. We have delimited subspace U_{b_1} with an emerging image.

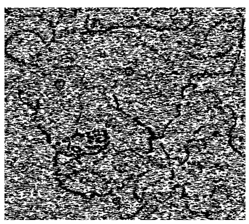

Figure 8: The ground space U_0, where the subspace U_{b_1} has been delimited with a system of nested emerging images (see lioness above)

Figure 9 shows a sample of the basis $\mathcal{B}$ of vector space U_0 used in the experiment. In this case, $\mathcal{B}$ has 196 elements $\big($i.e., $\sum\limits_{1\leq j\leq k} |B_{r_j}| = 196\big)$. It takes $1-5$ minutes to obtain an image with this basis in our scheme.

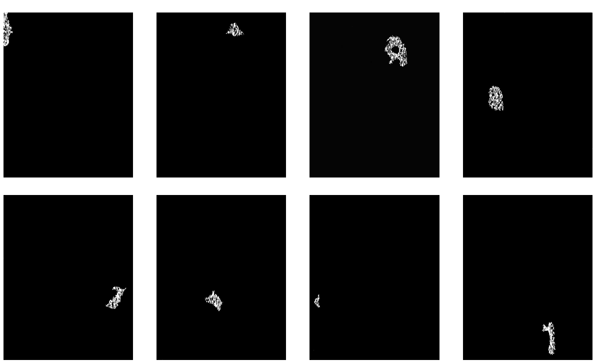

Figure 9: Sample of elements of the basis used in our algorithm

5 Conclusions and future work

We have used suitable linear representations of posets in order to define an automatic system for generating gestalts. Since previously proposed algorithm by Mitra et al generates only emerging images, the model described in this paper generalizes such scheme. Furthermore, a proper choice of the vector space U_0 and its corresponding fixed basis in a linear representation allows to decrease deformations produced by the nesting process.

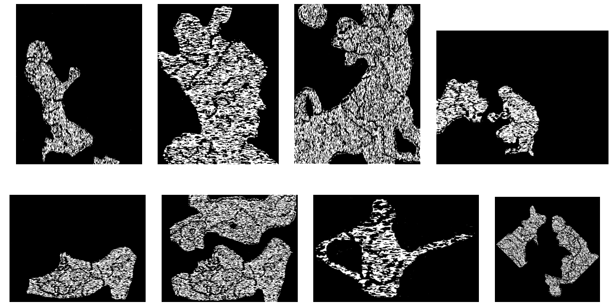

Figure 10: Some gestalts obtained by our algorithm. Images 6 and 7 (lion and lamp of Aladdin, respectively) are examples of ambiguities. Images 1, 2, 3, 4, 5, and 7 (from the left to the right) are part of a random sample of images, with the help of them, we conducted a poll in order to estimate the level of perception of images obtained for the proposed scheme

	Age			
Image	7 - 13	14 - 16	17 - 24	25 - 87
Woman kneeling in prayer	92%	95%	99%	96%
Queen	71%	82%	90%	91%
Dog	94%	93%	96%	99%
Feeding a dog	94%	95%	95%	97%
Shoe	85%	95%	90%	94%
Lamp	48%	66%	81%	84%

Table 1: Results of a poll of 2400 people conducted by the authors in order to determine the level of perception (lp) of a random sample of images obtained by the proposed algorithm.

Since in this paper, we have used only one basis of a fixed vector space in a linear representation to generate a large number of emerging images and ambiguities. It is interesting future work to define a category of linear representations of a poset whose objects generate an unlimited number of gestalts. In fact, we believe that the proposed scheme may be used to create an unlimited number of segmentation-resistant mechanisms for image-based CAPTCHAs.

Acknowledgements

The authors are grateful to the anonymous referees for their useful comments and suggestions.

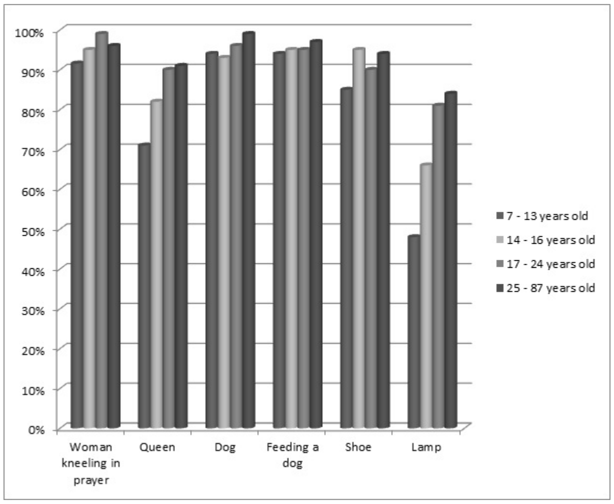

Figure 11: This diagram compares results given in Table 1 for images 1, 2, 3, 4, 5 and 7 in Figure 10. This study analyzed lp of images obtained by our algorithm. To do that, we picked 4 random samples of 600 people aged 7-13, 14-16, 17-24 and 25-87. These results allow to claim that lp of images generated by the proposed algorithm is very good.

References

[1] N. J. Mitra, H. K. Chu, T. Y. Lee, L. Wolf, H. Yeshurun, and D. Cohen-Or, *Emerging Images*, ACM Transactions on Graphics **28** (December, 2009), no. 5, 8 pp.

[2] A.M. Cañadas, *Morphisms in categories of representations of equipped posets*, JPANTA **25** (2012), no. 2, 145–176.

[3] ———, *Categorical properties of the algorithm of differentiation VII for equipped posets*, JPANTA **25** (2012), no. 2, 177–213.

[4] P. Gabriel, *Représentations indécomposables des ensemblés ordonnés*, Semin. P. Dubreil, 26 annee 1972/73, Algebre, Expose **13** (1973), 301–304.

[5] M. Hazewinkel, N. Gubareni, and V.V. Kirichenko, *Algebras, Rings and Modules*, First Edition, Vol. 2, Springer, 2007.

[6] M. Blum, L. von Ahn, and J. Langford, *Telling Humans and Computers Apart Automatically*, Communications of the ACM **2** (February 2004), no. 47, 57-60.

[7] C. Romero and E. Izquierdo, *A Survey of CAPTCHAs: Are Computers Getting the Better of us?*, Preprinted.

[8] J. Kim and F. Pellacini, *Jigsaw image mosaics*, Proc. SIGGRAPH (2002), 657–664.

[9] J. Tsotsos, *On the Relative Complexity of Active Vs. Passive Visual Search*, Int. Journal of Computer Vision **7** (1992), no. 2.

[10] R. Klette and P. Zamperoni, *Handbook of Image Processing Operators*, Second Edition, John Wiley & Sons, 1996.

[11] J. Russ, *The Image Processing Handbook*, Third Edition, CRC Press, 1998.

[12] L. Wolf, A. Zaritsky, and E. Rotman, *EI CAPTCHA-Emerging images CAPTCHA description page* (2010). www.cs.tau.ac.il/∼haimnach/eicaptcha/description/index.php.

[13] A. Desolneux, L. Moisan, and J. M. Morel, *From Gestalt Theory to Image Analysis*: *a probabilistic approach*, Springer, 2008.

[14] G. Kanizsa, *Grammatica del vedere / La Grammaire du voir*, Il Mulino Bologna / Éditions Diderot, arts et sciences, 1980/1997.

[15] S. Lehar, *Gestalt isomorphism and the primacy of subjective conscious experience: A Gestalt Bubble model*, Behavioral and Brain Sciences **26** (2003), 375-444.

[16] P. Moore and Ch. Fitz, *Using Gestalt Theory to Teach Document Design and Graphics*, Technical Communication Quarterly **2** (Fall 1993), no. 4, 389-410.

[17] A. M. Cañadas and N. P. P. Vanegas, *Extended Visual Cryptography Scheme with an Artificial Cocktail Party Effect*, Proc. 4th ICDP 2011, IEEEXplore, Digital Library (2011), 1-10. DOI: 10.1049/ic.2011.0114.

[18] Q. Tong, S.H. Zhang, R.R. Martin, and P.L. Rosin, *Nested images*, Preprint (2012).