

# Top-down visual search in Wimmelbild

Julia BERGBAUER

Technical University, Munich, GERMANY

Sibel TARI\*

Middle East Technical University, Ankara, TURKEY

## ABSTRACT

Wimmelbild which means “teeming figure picture” is a popular genre of visual puzzles. Abundant masses of small figures are brought together in complex arrangements to make one scene in a Wimmelbild. It is picture hunt game. We discuss what type of computations/processes could possibly underlie the solution of the discovery of figures that are hidden due to a distractive influence of the context. One thing for sure is that the processes are unlikely to be purely bottom-up. One possibility is to re-arrange parts and see what happens. As this idea is linked to creativity, there are abundant examples of unconventional part re-organization in modern art. A second possibility is to define what to look for. That is to formulate the search as a top-down process. We address top-down visual search in Wimmelbild with the help of diffuse distance and curvature coding fields.

**Keywords:** Camouflaged shapes, picture hunt, modern art, distractive influence of context, diffuse distance fields

## 1. TEEMING FIGURE PICTURES

Wimmelbild which means “teeming figure picture” is a popular genre of visual puzzles. Abundant masses of small figures are brought together in complex arrangements to make one scene in a Wimmelbild (Figure 1). It is picture hunt game. As such, it has the potential to be an excellent test bed for top-down visual search algorithms.

Figure hunting games starting with early Gestalt have played a noticeable role in visual perception studies. Gestalt psychologists used visual puzzles in which simpler shapes are hidden within more complex organizations to study the influence of experience on perception and the extent to which wholes influence the perception of parts. The earliest of these studies is the 1926 Gestalt experiments by Kurt Gottschald (Gottschald Hidden Figures Test) [1, 2]. Some examples are shown in Fig. 2 where each of the three complex organizations contains instances of a hexagon. Today, a quick search in internet retrieves wonderful children-oriented samples of the genre of Wimmelbild as well as striking samples of camouflage art, forming a wider genre of camouflaged objects pictures, e.g. by Octavio Ocampo [3]. These pictures vary in style, but hardly any of them seem amenable to purely bottom-up analysis. In some of them, it may be possible that processing starts bottom-up, delivering initial groupings as well as locations of salience, e.g. the hidden lion by D. Rust [4]. For some supposedly simpler black-and-white drawings, however, it appears to us that the process is highly likely to start top-down. For example, can you locate a hidden clover in Figure 3? Locating the clover is very difficult for a naked eye. It is hard to imagine any generic bottom-up bootstrapping process delivering initial grouping or salience regions. Generic parsing rules simply fail.

---

E-mail: [stari@metu.edu.tr](mailto:stari@metu.edu.tr) (Corresponding Author)

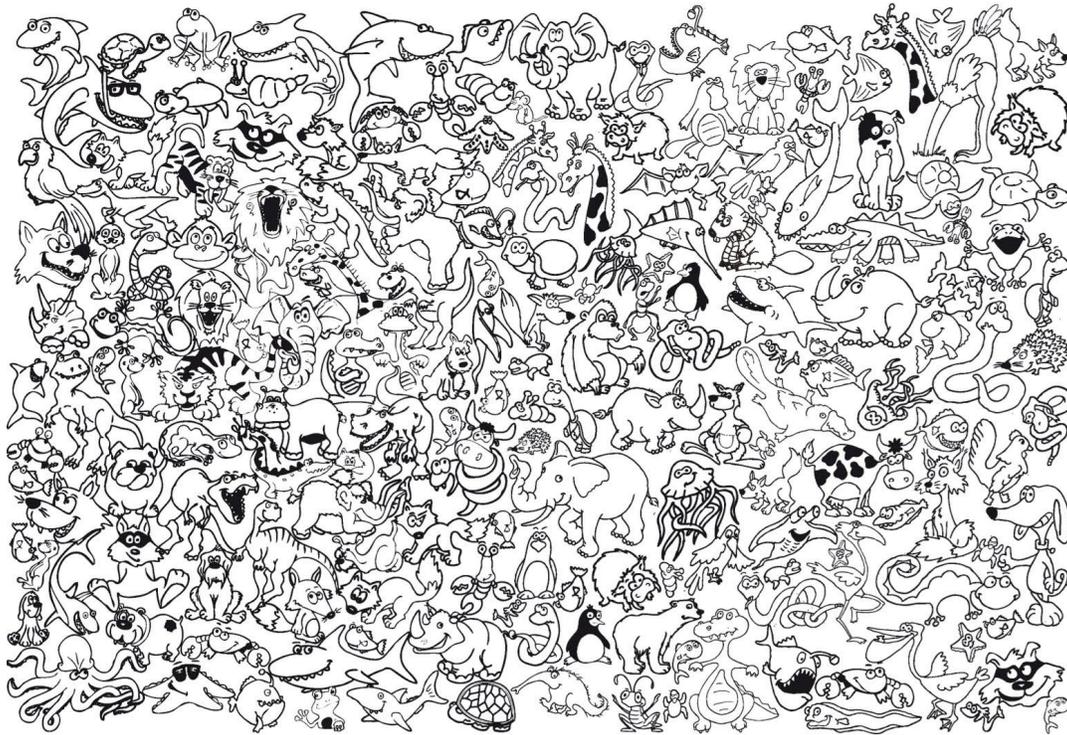


Figure 1. Zoo Wimmelbild (<http://corinnasteukmann.blogspot.com>)

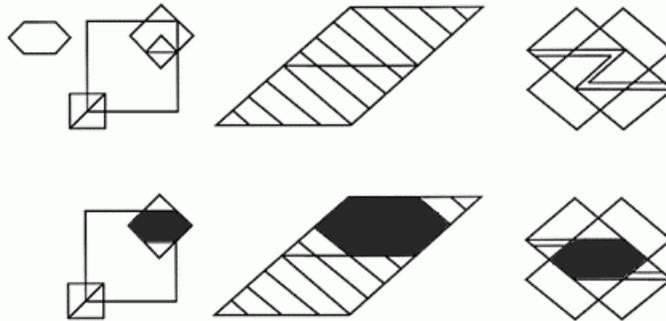


Figure 2. Gottschald Hidden Figures. Each of the three complex organizations contains instances of a hexagon.



Figure 3. A clover is hidden. It is hard to imagine any generic bottom-up bootstrapping process delivering initial grouping or salience regions. Generic parsing rules simply fail.

## 2. DISCOVERY OF HIDDEN FIGURES

What computations could possibly underlie the solution? How may one trace out different forms coexisting? For a moment, let us consider a much simpler drawing: a hexagram. The simplest way to draw a hexagram is to use two identical equilateral triangles. But once it is formed, seven constituent components (six triangles surrounding a hexagon) are easily perceived. Our perception switches between two different but equally valid parses. Yet there are other shapes hidden in the hexagram, e.g. parallelograms. Some parts are easily perceived and some are not, but all these shapes simultaneously exist in the whole. The hexagram is not a composition of independent components. It appears that the major source of difficulty is the non-compositional aspect of the picture. To discover a hidden figure, one has to undo the distracting influence of the context.

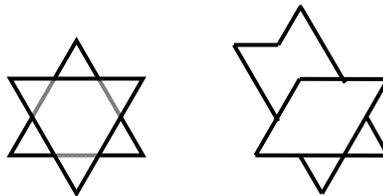


Figure 4. Interacting with the drawing makes hidden shapes explicit. A parallelogram in the hexagram becomes explicit.

### 2.1 Re-organization as a means of eliminating the distractive influence of context

Interestingly, an individual's dependence on or independence from the influence of distracting visual field has been considered as an indicator of his creativity because such a process involves both dividing wholes into parts (analysis) and organizing different part combinations (synthesis) without losing the sight of the total Gestalt. (This coincides with the definition of creativity by the Gestalt school.)

Thus, it is no surprise that there are abundant examples of this in art. A relatively recent attempt by Ursus Wehrli in re-organizing visual components in the paintings of modern painters may be worthy of mentioning [5]. Wehrli, in his book titled "Art tidied up" [6], re-arranges parts of paintings in a linear histogram like order. A naive eye may be fooled by the chosen title and the orderly appearance of his examples into thinking that what he is doing can be explained in terms of simple automatic clustering process based on well defined features such as size, color etc. The key point, however, seems something far off. Let us examine the re-organization of a Picasso painting by Wehrli, which is shown

in Figure 5. Clearly, it is the eye (Wehrli, in this case) that picks up the pieces and their new layout. This is not clustering! A joyful discussion on re-arranging parts in the general setting of creative design can be found in [7].



Figure 5. Picasso re-organized by Ursus Wehrli [4,5]. A naive eye may be fooled by some of his examples as well as his book's title into thinking that what Wehrli is doing can be explained in terms of simple automatic clustering process based on well defined features such as size, color etc. Clearly, it is the eye (Wehrli, in this case) that picks up the pieces and their new layout. This is not about clustering.

## 2.2 Top-down visual search as a means of eliminating the distractive influence of context

A solution more suitable for computer implementation is to let a specific target shape guide the search, i.e., to know what to look for. As computer vision scientists become more and more keen on evidence for top-down flow in visual processing, techniques for incorporating higher-level priors to lower-level processes become more and more widespread, delivering many interesting representation schemes and algorithms. In many cases, these techniques are shaped by applications, for example, anatomic template based segmentation has a wide applicability in medical imaging where the organ to be segmented is defined in the form of a simple closed curve defining its outline. These curves are either embedded as the zero-level curve of a function or defined via some parametric form such as splines. The search is cast as minimization of the deviation, hence maximization of the goodness-of-fit, between the template at a certain pose and scale and the picture at the respective locality. We treat the discovery of hidden shapes in teeming figure pictures in the same manner.

The main difference in our approach is how the shapes are represented. Firstly, our approach is non-parametric. Secondly, unlike level set based techniques, we do not treat shapes as entities sharply defined by a simple closed contour. Our shapes are *diffuse shapes*.

## 3. DIFFUSE SHAPES

The key idea is to diffuse the information localized on a contour to a plane in which the contour is embedded. An illustration is provided in Figure 6. The figure on the left depicts a clover, whereas the one on the right its diffuse representation. The information contained on the contour is diffused to a rectangular subset of the plane in which the contour is embedded. Once diffuse shapes and diffuse pictures are formed (Figure 7), the task gets reduced to locating a template in the picture. Diffuse representation is important to have smooth well-conditioned dependence of the goodness-of-fit to pose and scale parameters. A noteworthy remark is that without suggested field models the search does not work [8].

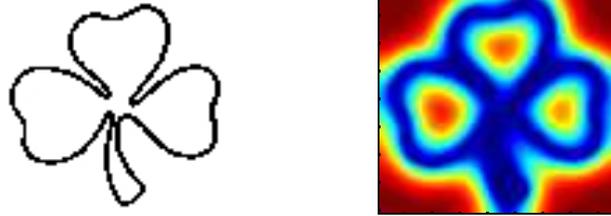


Figure 6. Diffuse shapes. (left) A clover. (right) A diffuse representation is given in pseudo-color.

Diffuse representations can be obtained in a variety of ways. The first thought could be to simply apply a low pass filter as in Gaussian scale space [9]. The propagation of the information, however, should be in such a way that propagated values reflect distances to the contour as well as local features such as curvature. An apparatus that will deliver the desired result is already available: the edge strength function in [10,11]. It is nothing but the solution of a screened Poisson PDE (partial differential equation). Issues related to scale and non-locality have been discussed in later works, respectively in [12] and [13]. In the edge strength function, the value at a point is approximately an exponentially decaying function of the minimum distance to the contour. The value also depends on the curvature at the nearest contour point. Thus, the edge strength function is a curvature aware distance field. Instead of using the edge strength function, we apply low pass filtering to a thickened form of drawing. In a future paper, we explain how this heuristic idea produces an approximation to the curvature aware edge strength function.

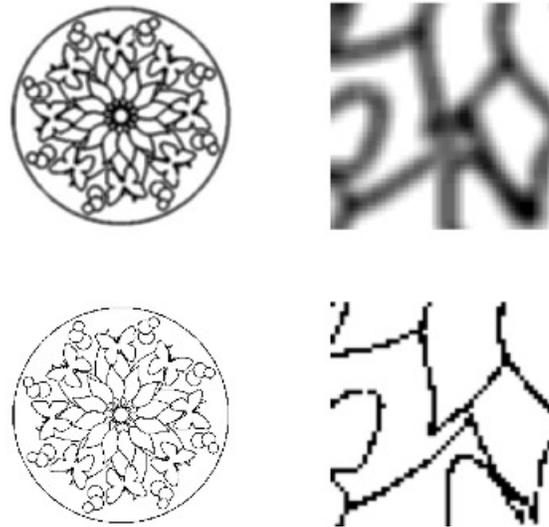


Figure 7. Diffuse forms for both the target shape and the whole picture should be formed. On the bottom left a Mandala pattern is shown. On the bottom right, an enlarged form of a piece cut out from the pattern is depicted. On the top row, diffuse representations of both the Mandala and the cut out piece are depicted.

In Figure 8, we depict 8 different poses of the target shape that we detect. Clearly, the suggested diffuse representation is insensitive to pose, as well as scale (Figure 9). Figure 9 illustrates the scale invariance of our scheme. We cut out a circle of a certain size from a Mandala-circles picture, and then search it back. Circles of varying sizes are retrieved.

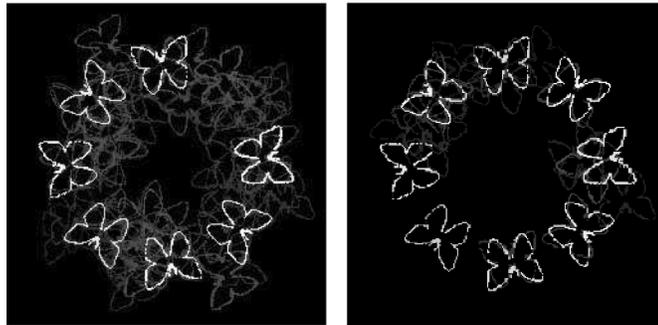


Figure 8. Illustrating the insensitivity of the diffuse representation to pose changes. Several hypotheses are generated. Brightness is a measure of goodness-of-fit. Different poses of the cut out butterfly piece are retrieved, each with similar goodness-of-fit. On the left independent search experiments are averaged.

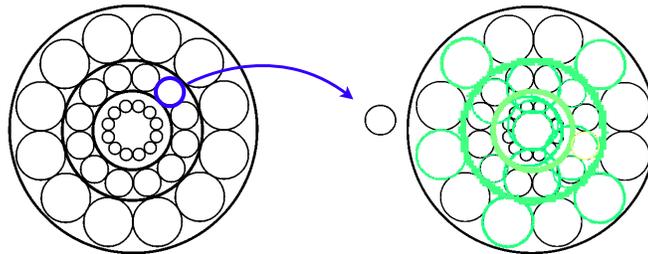


Figure 9. Scaling. Circles of varying scales in the Mandala pattern can be detected. The blue circle is used as the probe shape.

#### 4. EXPERIMENTS

We present the results of our experiments on the two teeming figure pictures. In the first search experiment, we search for the elephant in the Zoo Wimmelbild. In the second search experiment, we search for the clover in the picture shown in Figure 3. For the search, we use nondeterministic optimization algorithm, i.e., genetic algorithm, so that we can readily generate multiple hypothesis.

In Figure 10, the result of the search for the elephant (top row) in the Zoo is depicted. Each generated hypothesis is tone coded based on the goodness-of-fit (bottom right). The correct fit has significantly higher goodness-of-fit than all the others. The best hypothesis is shown in red in the bottom left figure.

In Figure 11, the result of the search for the clover. Again several hypotheses are generated; the one with the highest goodness-of-fit is the correct one.

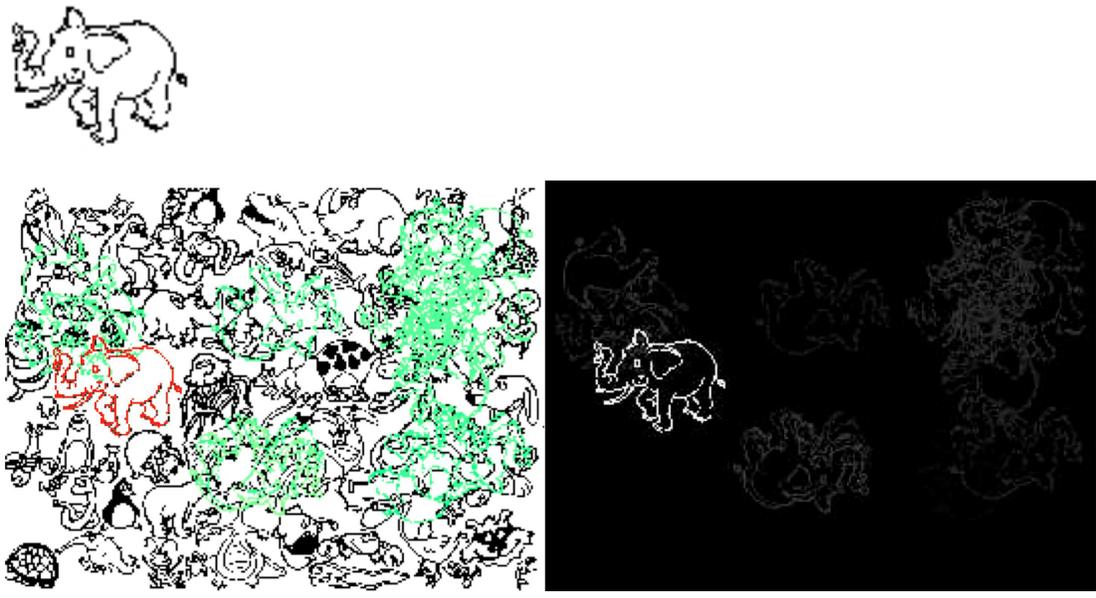


Figure 10. Hunt for the elephant.

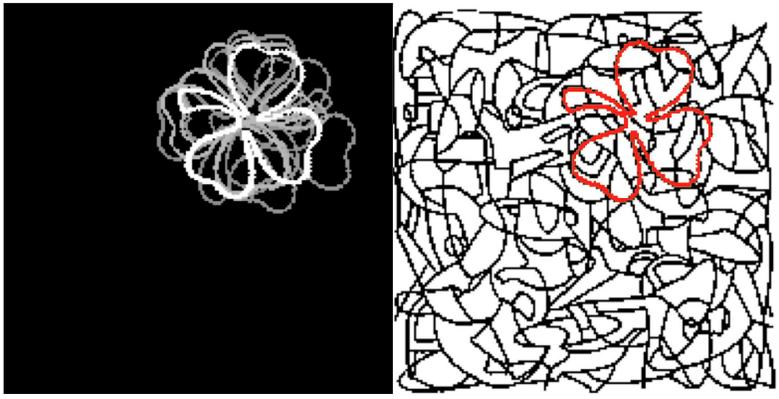


Figure 11. Hunt for the clover. The strongest hypothesis is shown in red on the picture.

## CREDITS

S.T. acknowledges the generous support of the Alexander von Humboldt Foundation. The work is also funded by TUBITAK under the grants 108E015 and 112E208. The work of J.B has been completed as a partial fulfillment of the course project supervised by S.T. while she was a guest professor at TU Munich Department of Mathematics. The author thanks Dr. Ulrike Albrecht of the Alexander von Humboldt Foundation for pointing out the relevance of the work of Ursus Wehrli [5].

## REFERENCES

1. Gottschald, K., "The influence of past experience on the perception of figures," [Experiments in Visual Perception], Penguin Books, Baltimore, MD (1966).
2. Gottschald, K., "Über den Einfluß der Erfahrung auf die Wahrnehmung von Figuren. I. Über den Einfluß gehäufte. Einprägung von Figuren auf ihre Sichtbarkeit in umfassenden Konfigurationen," *Psychologische Forschung* 8, 261-317 (1926).
3. Octavio Ocampo Pictures, <<http://octavioocampo.com.mx/>> (17 January 2013). [octavioocampo.com.mx/](http://octavioocampo.com.mx/)
4. Rust, D., "Hidden Lion", <<http://www.planetperplex.com/en/item/hidden-lion/>> (15 January 2013). [www.planetperplex.com/en/item/hidden-lion/](http://www.planetperplex.com/en/item/hidden-lion/)
5. Albrecht, Ulrike. *Personal correspondance to S. Tari* on Nov 2010.
6. Wehrli, U., [Kunst Aufraeumen], Kein & Aber, Zürich (2002).
7. Stiny, G., [Shape: Talking about seeing and doing], MIT Press, Cambridge, MA (2006).
8. Keles, H., Ozkar, M. and Tari, S., "Weighted shapes for embedding percpetual wholes," *Environment and Planning B: Planning and Design* 39(2), 360-375 (2012).
9. Koenderink, J., "The structure of images", *Biological Cybernetics* 50, 363-370 (1984).
10. Tari, S. and Shah, J., "Local Symmetries in Arbitrary Dimensions," *IEEE International Conference on Computer Vision (ICCV)*, (1998).
11. Tari, Z.S.G., Shah, J. and Pien, H., "Extraction of shape skeletons from grayscale images," *Computer Vision and Image Understanding* 66, 133-146 (1997).
12. Aslan, C. and Tari, S., "An Axis Based Representation," *IEEE International Conference on Computer Vision (ICCV)*, (2005).
13. Tari, S. and Genctav, M., "From a modified Ambrosio-Tortorelli to a randomized part hierarchy tree," *SSVM*, Springer-Verlag, 267-278 (2011).