

Wimmelbild Analysis with Approximate Curvature Coding Distance Images

Julia Bergbauer¹ and Sibel Tari²

¹ TU Munich, Garching bei Muenchen, DE-85747

² Middle East Technical University, Ankara, TR-06800
stari@metu.edu.tr**

Abstract. We consider a task of tracing out target figures hidden in teeming figure pictures come to known as *Wimmelbild(er)*. Wimmelbild is a popular genre of visual puzzles; a timeless classic for children, artists and cognitive scientists. Particularly suited to the considered task, we propose a diffuse representation which serves as a heuristic approximation mimicking curvature coding distance images. Curvature coding distance images received increased attention in recent years. Typically, they are computed as solutions to variants of Poisson PDE. The proposed approximation is based on erosion of the white space (background) followed by isotropic averaging, hence, does not require solving a PDE.

Keywords: Poisson PDE and its variants; level sets; non-linear diffusion; figure-hunt games; teeming figure pictures; applications of variational and PDE methods

1 Introduction

Level set methods have been successfully applied to knowledge based segmentation; bringing in an ability to deal with physically corrupted incomplete data. Most typically, shape knowledge is coded via signed distance transform, embedding the $1 - D$ shape boundary as the zero-level set of a function defined on a connected bounded open subset of \mathbb{R}^2 [9, 8]. It is also possible to replace the sharp interface model in level set based segmentation methods with diffuse ones [12, 10], decaying exponentially and coding curvature.

In recent years, there is a growing interest in exponentially decaying curvature coding distance images with examples including [1, 5, 2, 11, 12].

The idea is to replace a point set S denoting possibly incomplete object boundaries with a smooth function $\nu : R \rightarrow \mathbb{R}$ where $R \subset \mathbb{R}^2$ is an open connected bounded set s.t. $R \supset S$. The function ν is the solution of

$$\nabla \cdot (\nabla \nu) - \alpha^2 \nu = 0 \tag{1}$$

subject to boundary conditions:

$$\nu \Big|_S = 1 \quad \text{and} \quad \partial_\eta \nu \Big|_{\partial R} = 0 \tag{2a,b}$$

** Corresponding author

where ∂R denotes the boundary of R and $\partial_\eta \nu$ the derivative of ν in the normal direction to ∂R . Interestingly, from [12],

$$\nu(x, y) \approx \frac{1}{2\alpha^2} \left(2\alpha + \text{curv}(x, y) \right) \frac{\partial \nu}{\partial \eta} + O\left(\frac{1}{\alpha^3}\right) \quad (3)$$

where $\text{curv}(x, y)$ is the curvature of the level curve passing through the point (x, y) at (x, y) . Eq. (3) indicates a reciprocal relationship between the level curve curvature and the gradient of ν . That is, unlike the usual distance transform, ν is an implicit coder of the curvature, a valuable geometric feature, without explicit estimation of higher order derivatives. (One of the original goals in proposing ν was to bridge low level and high level vision [11, 12]).

In this paper, in the setting of a specific task, namely tracing out target figures hidden in teeming figure pictures, we present a much simplistic way of obtaining an analogous (curvature dependent) behavior in a band around S . Our computation does not require the computation of the entire ν function on the entire domain R by solving a PDE.

Preliminary experiments are highly encouraging and indicate the potential of the approach.

2 The problem setting

Wimmelbild is a popular genre of visual puzzles. It means teeming figure picture. Abundant masses of small figures are brought together in complex arrangements to make one scene in a *Wimmelbild*, to be used for a figure hunt game (Fig. 1).

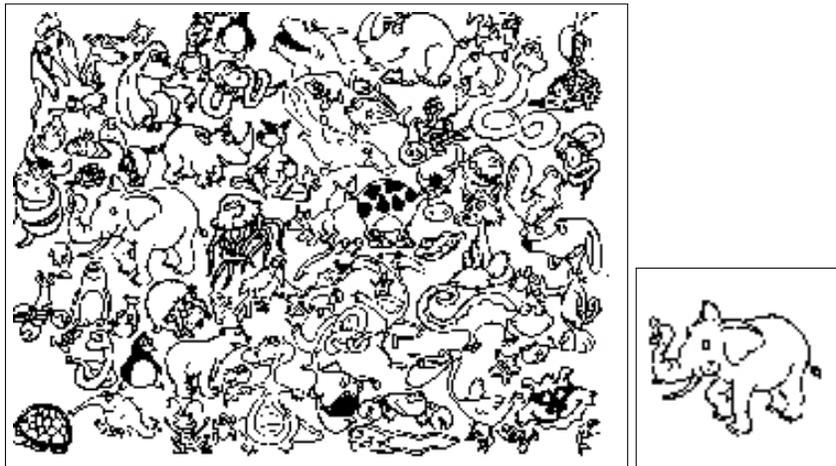


Fig. 1. An elephant in a zoo. A sample wimmelbild (left) and an elephant figure (right) to hunt for in the Zoo Wimmelbild.

Figure-hunt games have been a timeless classic for children, artists and cognitive scientists. As early as 1926 Kurt Gottschald experimented with intentionally designed hidden figures – simple drawings where simple shapes such as polygons are embedded within more complex organizations– to study the influence of experience on perception and the extent to which wholes influence the perception of parts [4].

There are of course colorful versions of these genre of visual puzzles. Interestingly, sketch-like black and white versions of teeming figure pictures pose even a bigger challenge than their colored counterparts:

Firstly, during a hunt for the elephant, suppose we somehow landed on the correct location in the picture yet hypothesized a wrong scale, say a smaller scale, such that the hypothesized elephant fits inside the white space (background) surrounded by the figural loci of the actual elephant, yielding a matching cost of zero; providing, therefore, no clue as to whether we are in the right neighborhood, on a white space, or on a region containing a figure which has no common elements with the elephant. Whereas each point on a dense picture (be it color or gray) is informative, information in sketch-like binary pictures is concentrated on loci of lower dimension, *i.e.*, curves denoting figural loci.

Secondly, as a consequence of lower dimensionality of the figural loci, final pictures could get extremely complicated. Observe that it is impossible to trace out the hidden clover in Fig. 2 using Gestalt parsing rules. The final picture is not simply a superposition of figures. Indeed, figures are first added, but then thresholded.

Therefore, the task can not be simply cast as locating a subpicture within a whole picture.

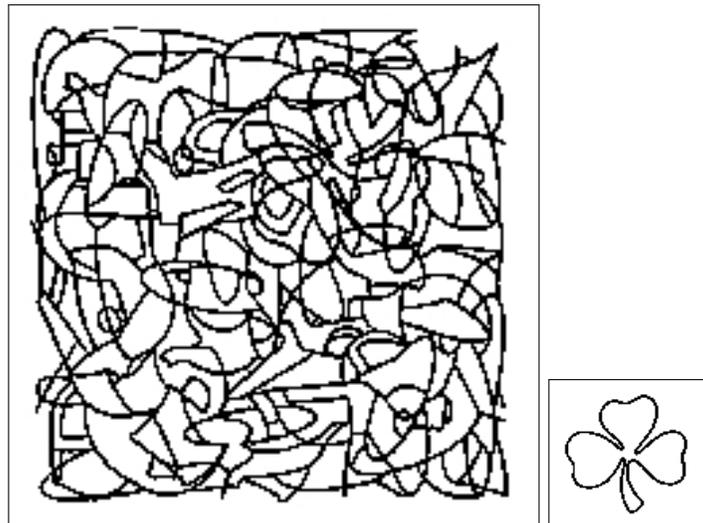


Fig. 2. Can you trace out the hidden clover?

Related Work To the best of our knowledge, *Wimmelbild(er)* has not been studied within the variational and PDE methods community before. Saarbruecken group recently presented an inpainting based steganography application [7]. The nature of their problem hence their solution strategy, however, are quite different. Their goal is to hide a secret image by embedding it into arbitrary cover images. Both the secret and the cover are dense images and recovery of the secret is possible only via a password. Therefore, ordinary observer can not detect whether an image contains a secret or not. In reconstructing frescos, Fornasier et.al. [3] addressed the problem of locating small fragments within a whole; for each small piece of plaster that still showed an element of the design of the fresco, they were able to find where it belonged. Such approaches can not be used due to non-additive and non-linear nature of the binary drawings that we consider (such as Fig. 2). Curvature coding distance fields are becoming more and more commonplace. Theoretical investigations can be found in [12, 1, 5, 2], among others. They have been used to address a variety of shape related problems, including skeleton computation and knowledge based segmentation as well as other vision problems.

3 Our Approach

The key idea is to propagate information restricted to figural loci to neighboring areas so that it becomes possible to know whether a location is close or far away from the desired locations.

We start by uniformly eroding the white space, or equivalently, dilating the figure. Hence, the drawings (both the *wimmelbild* and the target figure) become thicker. Then, we diffuse by computing a local isotropic average. It is sufficient to compute the local average only for the points falling on the thickened figural loci or in a slightly wider band surrounding it. This transforms the sketch-like binary drawing to a gray-tone picture which may be referred as a diffuse drawing. This diffuse drawing is an approximation to a curvature coding distance image.

The rightmost column in Fig. 3 depicts the outcome of the described procedure. The leftmost column is the original drawing and the fourth column from the left is the thickened drawing. If the averaging and the dilation radii are identical, the highest value is attained on the figural loci; from thereof values decrease as a function of distance in the normal direction. Thus, diffusion produces iso-intensity contours, each following the figural loci from a fixed distance. The lower the intensity, the further away the iso-intensity curve from the figural loci. The two columns in the middle (second and third from the left) depict the results of two different local isotropic averaging applied to the original thin drawing. There, one can not observe the distance-coding behaviour, *i.e.*, the initial thickening is a crucial step.

What makes the iso-intensity contours of our diffuse drawings further interesting is that they implicitly code curvature: Let us select two locations on the white space of the drawing such that the nearest figural point to the first location has a high curvature, while on the contrary, the nearest figural point to the

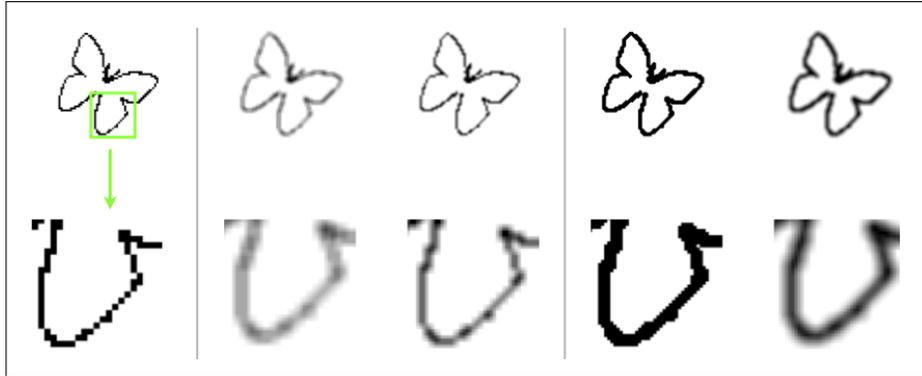


Fig. 3. Mimicking curvature coding distance via erosion of the white space (dilation of the figural loci) followed by averaging.

second location is on a flat part. Let us make sure, however, that their distances to their respective nearest figure points are equal. That is each of the points have the same normal distance to the drawing. For example, let us consider the two locations marked by the intersection points of the two crosses in Fig. 4.

Note that even though the normal distances from each location to the figural loci are identical, the distances in the tangential directions are not.

In the tangential direction, indeed, the first location (marked by the cross on the upper right in Fig. 4) is closer to the figural loci, causing the average value at that location to be higher compared to the average value at the second location (marked by the cross on the lower left in Fig. 4). Consequently, the level curve passing through the second location will not pass through the first location but through a location further down in the direction of inward normal.

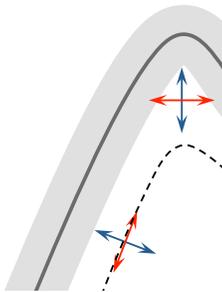


Fig. 4. Two locations of identical normal distance to the figural loci are marked by the two crosses. In the tangential direction, however, the location of which nearest figural point has a higher curvature is closer to figural loci.

As a result, within a band surrounding the figural loci, our diffuse drawing (obtained by dilation followed by isotropic diffusion) mimics a curvature coding distance field similar to the ν , the solution of a damped Poisson PDE [12]. In Fig. 5, we compare our diffuse drawing to a usual distance image (i.e. Eikonal PDE with constant right hand side) restricted to a band surrounding the figural loci. Whereas the effects of discretization noise remains (even amplified) in the usual distance image, the iso-intensity contours in our diffuse model *smoothly* follow the boundary.

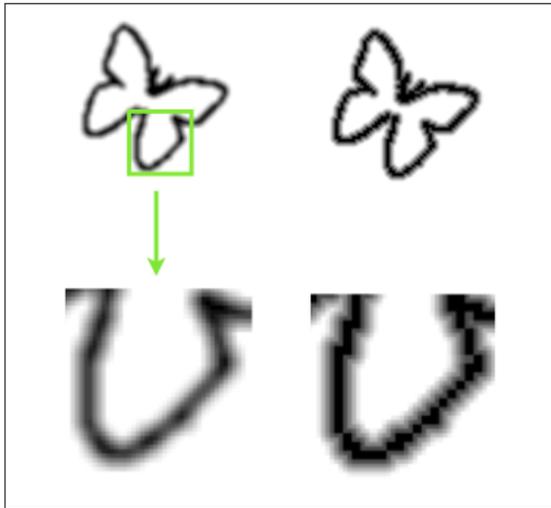


Fig. 5. Curvature coding distance image obtained by thickening + isotropic averaging (left) versus the usual distance image (right).

We avoid solving Poisson PDE or variants for two reasons. Firstly, our approximation is both easier and faster to compute. But more importantly, a Poisson based distance field, being the steady state solution to a biased diffusion equation, $\frac{\partial \nu}{\partial \tau} = \nabla \cdot (\nabla \nu) - \alpha^2 \nu$, is too much influenced by long-range interactions among opposing boundaries. This may be detrimental if several figural loci overlap as in Fig. 2.

Once the drawings (both the wimmelbild and the target figure to be hunted for) are converted to diffused forms, the best match is formulated as finding the deformation parameters (e.g. scale, location and orientation) that yield the best match. The matching cost is measured as the sum of the gray value differences between the wimmelbild and the target figure. Of course, the sum is taken over those locations that fall within the band surrounding the figural loci within which the diffuse field has been constructed. Moreover, the cost is normalized by dividing it to the number of locations contributed to its computation.

In Fig. 6, the cost calculation is illustrated. Observe how matching cost becomes informative when raw binary figures are replaced with diffuse ones.

Once the matching cost is defined, the optimizing parameters are determined via a probabilistic algorithm which returns multiple solutions. The importance of using such an algorithm is discussed in [6]. We use genetic algorithms based optimization which is readily available in Matlab environment. It minimizes an energy functional by varying its input variables. It is called via the comment:

```
[variables, energy] = ga(fitnessfcn, nvars, [], [], [], [], lb, ub, [], IntCon);
fitnessfcn = @energyfunctional, pattern, shape, diffusingparam;
```

The *output* includes the determined value of the vector of *variables* for rotation, scaling and translation as well as the value of the corresponding minimal *energy*. As *input* it requires a **fitnessfcn**. This has to be a functional which takes the *variables* as the first input. Its single output has to be the value of the corresponding energy. Furthermore, the parameters *nvars* (number of variables), *lb* (lower bound for variables), *ub* (upper bound for variables) and *IntCon* (an interval containing the indices of variables that should be integers) have to be specified.

Each call to **ga** returns the best fit obtained after a certain number of trials. Due to randomization, multiple calls to **ga** generate multiple hypotheses. We depict all the generated hypotheses after an ordering based on the matching cost.

4 Experimental Results

In this section, we will present our experiments with the two wimmelbilder:

- hunt for the elephant in the Zoo Wimmelbild;
- hunt for the clover in a heap of overlapping contours.

Prior to that, however, we test our approach on simpler drawings of repeated patterns: Mandalas. One purpose of these supporting illustrations is to examine the robustness of the method to scale and pose variations. The second purpose is to observe whether the genetic algorithm returns correct fits more often than the wrong ones.

First, to observe the robustness with respect to scale, we consider a composition of circles of varying size. One of the circles (shown in blue color on the left column of Fig. 7) is selected as the target figure. On the right column, we depict all the circles detected after several runs of the genetic algorithm. Observe that the method can handle scale variations.

Second, to observe the robustness with respect to pose, we consider a simple Mandala pattern (Fig. 8). The two subfigures extracted from it (on the right) are to be used as targets. We expect to find 8 instances of the butterfly target and 4 instances of the second target.

The results are depicted in the left column of Fig. 9. The top row depicts the results for the *butterfly* target and the second row for the second target. The

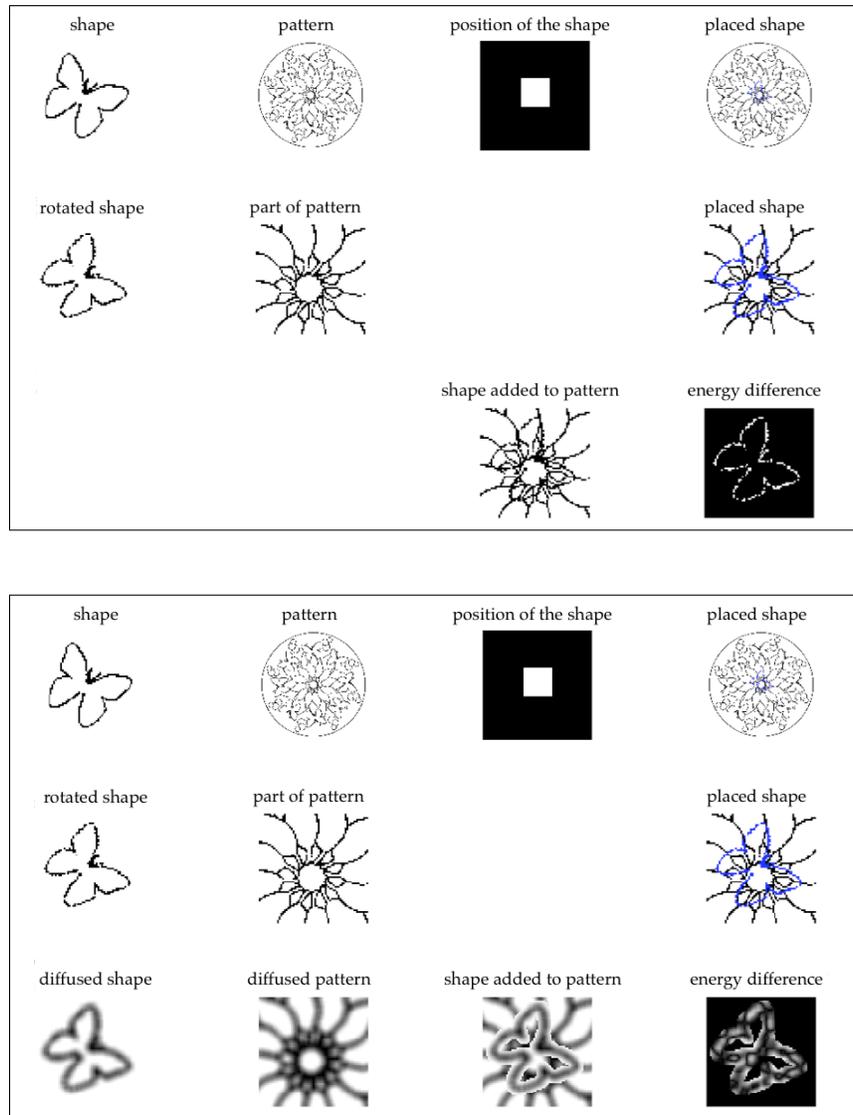


Fig. 6. Illustration of cost calculation. When raw binary drawings are used (top), it is hard to tell how well the target figure is located. Observe how matching cost becomes informative when raw binary figures are replaced with diffuse ones (bottom).

ga is invoked around 100 times and all of the 100 results are displayed. Each outcome of the **ga** is tone-coded based on the matching cost. Lighter the tone, lower the matching cost hence better the fit. The correct fits are found regardless

of their pose. Moreover, the matching cost is significantly lower for the correct fits. This indicates the robustness of the representation to pose changes.

Third, we have tested whether the good fits (those of lower matching cost) are obtained more often than the bad fits. This is important as the algorithm is not a deterministic one. We have performed independent **ga** runs, each run producing several hypotheses. We have then computed the average of the batches of independent runs. As the results shown in the right column compellingly demonstrate, **ga** has a tendency to return good fits more often than the bad ones.

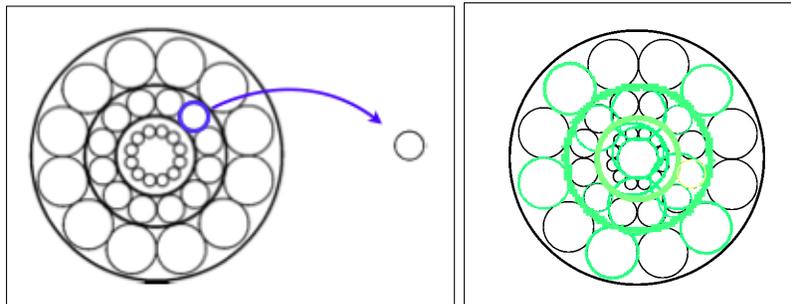


Fig. 7. Circles of varying scale.

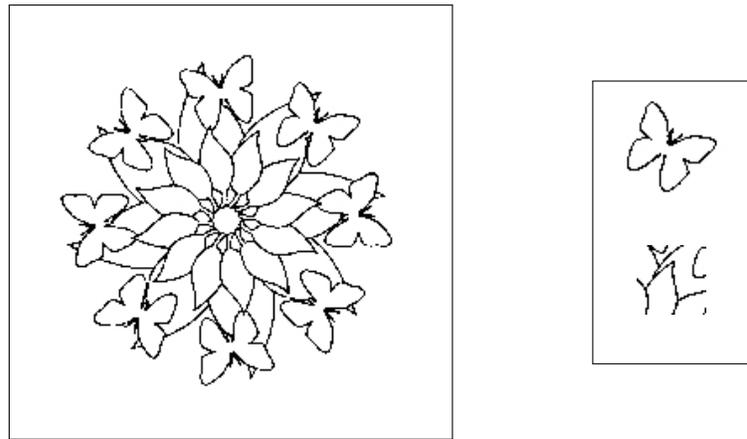


Fig. 8. A Mandala pattern and two subfigures extracted from it.

Finally, we present our two figure hunt experiments. The results of the elephant hunt is given in Fig. 10. On the left, multiple hypotheses produced by **ga**

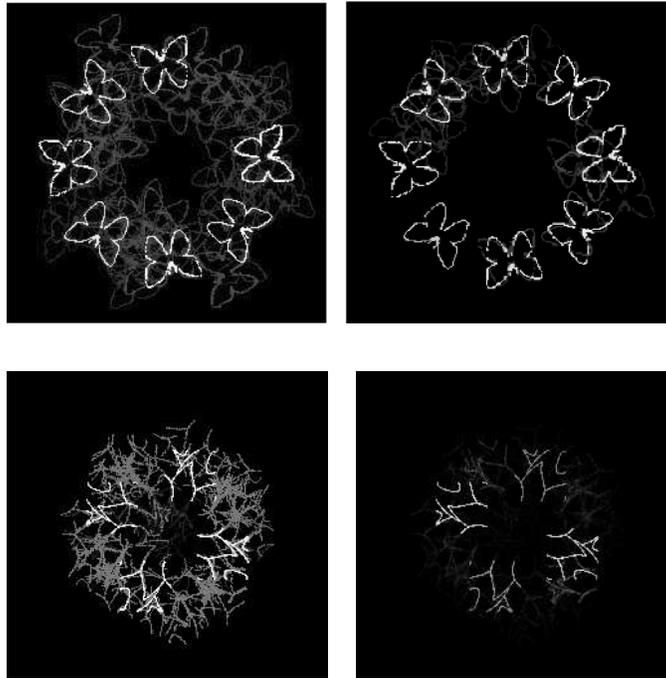


Fig. 9. (Left) several hypotheses returned by the genetic algorithm. Gray-tones reflect the matching cost. The lighter the tone, the better the fit.

are tone-mapped. The correct hypothesis is significantly lighter than all the others. On the right, the best match is shown in red, whereas the others in shades of green.

Fig. 11 depicts the results of the clover hunt. The generated hypotheses are depicted on the top left. The hypothesis with the lowest matching cost is depicted on the top right. The best hypothesis is traced with red marker on the original drawing (bottom left), and the original drawing is repeated (bottom right) as a convenience to the reader.

5 Summary and Conclusion

We have addressed the task of tracing out target figures in sketch-like binary teeming figure pictures. Particularly suited to the task, we propose a simple heuristic for generating diffuse drawings that imitate curvature coding distance images which are typically computed as solutions to elliptic PDEs. Our work extends the applications of diffusion based ideas to an interesting problem.

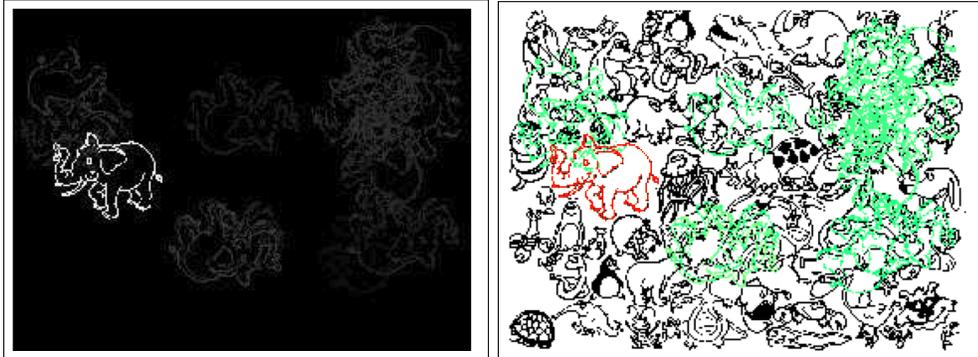


Fig. 10. Hunt for the elephant.

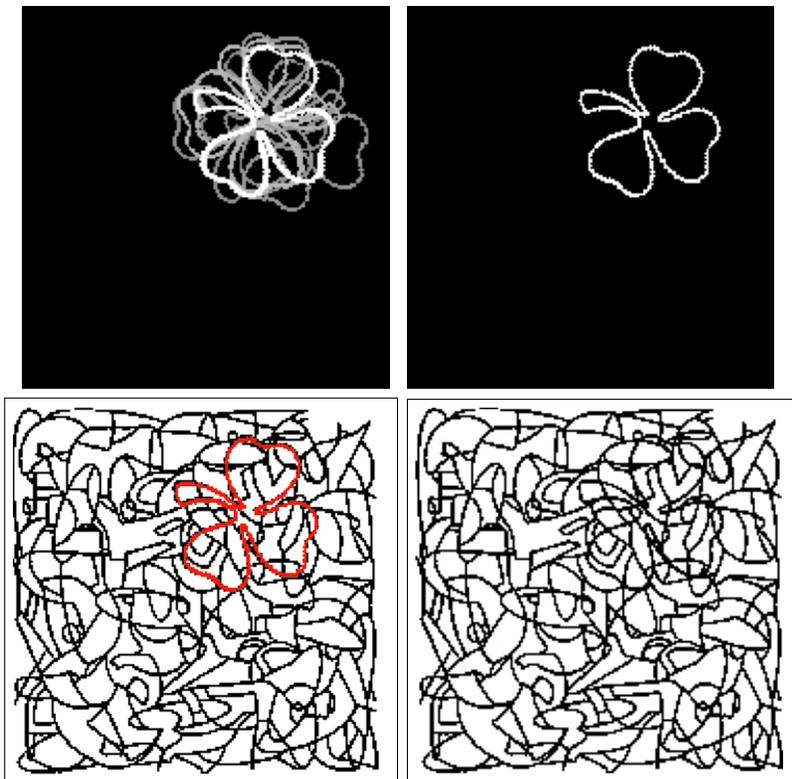


Fig. 11. Hunt for the clover.

Acknowledgements

The work of JB has been completed in March 2012 as a part of MA5313 course offered by ST, while ST was spending her sabbatical at Folkmar Bornemann Group at TU Munich. At present, JB is affiliated with D. Cremers Group at the same institution. Subsequent work in writing the paper has been completed at the Middle East Technical University.

ST thanks to Folkmar Bornemann for providing a truly scientific environment and flawless support during the time of her stay, and extends her gratitude to the Alexander von Humboldt Foundation for generous financial support.

References

1. G. Aubert and J.F. Aujol. Poisson skeleton revisited: A new mathematical perspective. *J. Math. Imaging Vis.*, 2012.
2. P. Dimitrov, M. Lawlor, and S.W. Zucker. Distance images and intermediate-level vision. In A Bronstein A Bruckstein, B ter Haar Romeny and M Bronstein, editors, *SSVM*, pages 653–664. Springer-Verlag, 2011.
3. M. Fornasier and D. Toniolo. Fast, robust and efficient 2d pattern recognition for re-assembling fragmented images. *Pattern Recognition*, 38(11):2074 – 2087, 2005.
4. K. Gottschald. Ueber den Einfluss der Erfahrung auf die Wahrnehmung von Figuren. *Psychologische Forschung*, 8:261–317, 1926.
5. K. Gurumoorthy and A. Rangarajan. A Schroedinger equation for the fast computation of approximate Euclidean distance functions. In M Lysaker XC Tai, K Morken and KA Lie, editors, *SSVM*, pages 100–111. Springer-Verlag, 2009.
6. H.Y. Keles, M. Ozkar, and S. Tari. Weighted shapes for embedding perceived wholes. *Environment and Planning B: Planning and Design*, 39:360–375, 2012.
7. M. Mainberger, C. Schmaltz, M. Berg, J. Weickert, and M. Backes. Diffusion-based image compression in steganography. In *ISVC (2)*, pages 219–228, 2012.
8. S Osher and N Paragios. *Geometric Level Set Methods in Imaging, Vision, and Graphics*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2003.
9. Nikos Paragios. A variational approach for the segmentation of the left ventricle in cardiac image analysis. *International Journal of Computer Vision*, 50(3):345–362, 2002.
10. H. Pien, M. Desai, and J. Shah. Segmentation of MR images using curve evolution and prior information. *Int J Pattern Recogn*, 11(8):1233–1245, 1997.
11. S. Tari and M. Genctav. From a modified Ambrosio-Tortorelli to a randomized part hierarchy tree. In A Bronstein A Bruckstein, B ter Haar Romeny and M Bronstein, editors, *SSVM*, pages 267–278. Springer-Verlag, 2011.
12. ZSG. Tari, J. Shah, and H. Pien. A computationally efficient shape analysis via level sets. In *Proceedings of the Workshop on Mathematical Methods in Biomedical Image Analysis*, pages 234–243, 1996.