EDGE POTENTIAL FUNCTIONS AND GENETIC ALGORITHMS FOR SHAPE-BASED IMAGE RETRIEVAL

Minh-Son Dao, Francesco G.B. De Natale, IEEE Member, Andrea Massa, IEEE Member

DIT – University of Trento
Via Sommarive, 14 – 38050 Trento, Italy

ABSTRACT

In this paper, a new approach to the image retrieval problem is presented, that uses Edge Potential Functions (EPF) and genetic algorithms (GA). The method allows a user to draw a rough sketch of the shape and to find or rank the images in a database that contain a similar shape at any position, rotation and scaling factor. It will be explained how GAs allow to exploit the capability of EPFs to attract a sketch contour: as a result, the algorithm provides the set of geometrical transformations corresponding to the best match, and a confidence factor about the presence of a matching object. The method has been widely tested achieving very satisfactory results.

1. INTRODUCTION

The great diffusion of Internet and Web-based applications and archives, has engendered a huge amount of distributed multimedia information, accessible through data networks. In this framework, the problem of content-based retrieval is gaining a high importance [1]. Among different techniques to perform a visual query, the possibility of using object shapes is very interesting, for it allows the user to search for a particular visual object by sketching the relevant shape on a simple graphical interface. On the other hand, this application is very demanding in terms of implementation constraints: in fact, browsing tools should be characterized by a nearly real-time response, low processing requirements, and high robustness.

In the literature, there are examples of techniques using mathematical shape models and optimization methods to perform the matching operation. In [2], the authors present a matching procedure based on the elastic deformation of the sketches over object shapes present in the target image; in this case, the similarity measure is in inverse relationship with the energy spent in the deformation the original sketch. In [3], a technique for matching binary shapes is proposed that uses affine transforms and genetic algorithms, applied to a polygonal approximation of the contour. In [4], the authors proposed the use of straight-line segments and a genetic algorithm to detect a 2D target model in cluttered environments and determine its rotation, scaling and translation.

The algorithm proposed in [5] uses GAs to detect partially occluded objects in a scene at any rotation, scaling and translation, by introducing the concept of prominent boundary fragments, defined as the sequences of significant contour points present on both reference and target objects.

In this paper, an innovative approach to shape matching is proposed, which is based on the concept of Edge Potential Functions. EPFs can be easily calculated starting from the edge map extracted from a digital image, and represent a sort of attraction field in analogy with the field generated by a charged element. In the context of shape matching, the EPF can be used to attract a user sketch in the position where a similar shape is present in the image: in fact, the higher the similarity of the two shapes, the higher the total attraction engendered by the edge field. In the proposed method, the matching algorithm combines EPF with a genetic algorithm to provide a powerful and reliable tool for edge-based shape matching.

In Sect. 2, the concept of EPFs is outlined; in Sect. 3, the procedure for finding a shape inside a digital image using EPFs is described, making use of an optimisation approach based on genetic algorithms (GA); finally, in Sect. 4, results are presented.

2. EDGE POTENTIAL FUNCTIONS

The basic concept of edge potential functions derives from the potential generated by charged particles. It is well known that a set of point charges Q_i in a homogeneous background generates a potential, the intensity of which depends on the distance from the charges and on the electrical permettivity of the medium ε, namely:

$$v(\vec{r}) = \frac{1}{4\pi\varepsilon}\sum_i \frac{Q_i}{|\vec{r} - \vec{r}_i|} \qquad (1)$$

where $\vec{r}$ *and* $\vec{r}_i$ are the observation point and charge locations, respectively.

Now, let us consider a test object with opposite charge immersed in the potential field: it would be attracted to the field point where the differential potential is maximized. In complete analogy with the above behavior, in our model, the i-th edge point in the image at coordinates (x_i, y_i) can be assumed to be equivalent to a point charge $Q_{eq}(x_i, y_i)$, contributing to the potential of all image pixels:

$$EPF(x,y) = \frac{1}{4\pi\varepsilon_{eq}} \sum_i \frac{Q_{eq}(x_i, y_i)}{\sqrt{(x-x_i)^2 + (y-y_i)^2}} \quad (2)$$

where ε_{eq} is a constant measuring the equivalent permittivity of the image background.

To complete the model, a generic sketch contour to be matched with the image content, can be considered as a test object in the equivalent edge potential field. Consequently, it is expected to be attracted by a set of equivalent charged points that maximizes the potential along the edge.

As far as binary edge maps are considered, Eq. 2 can be simplified by modeling each edge points as an equal charged element of value $Q_{eq}(x_i, y_i)=Q=1$, thus obtaining the binary EPF (BEPF):

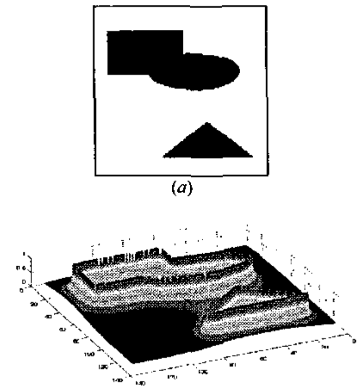
$$BEPF(x,y) = \frac{1}{4\pi\varepsilon_{eq}} \sum_i \frac{1}{\sqrt{(x-x_i)^2 + (y-y_i)^2}} \quad (3)$$

where ε_{eq} influences the spread of the potential function making it more steep or smooth depending on its large or small value. This property will be very important to manage the attractivity of the contours during the matching process, as explained in Sect. 3.

The above definition of BEPF presents two problems: first, singularity points occur at edge pixel locations (x_i, y_i); and second, the maximum of the potential field depends on the equivalent permittivity. In order to avoid such drawbacks, BEPF is clipped and normalized. Fig. 1 shows the BEPF obtained from a simple binary shape.

From the viewpoint of computational load, Eq. 3 implies a complexity of the EPF estimation proportional to the number of edge pixels present in the image. Another important drawback of EPFs lies in the fact the potential in a given image point is not only proportional to the distance from an edge, but also to the average activity of the image area. In particular, in an image area characterized by a high density of edges (e.g., a texture), the potential function is averagely high, independently of the distance from significant objects. A more effective computation procedure of EPFs that ensures a significant speed-up and provides a simple solution to the problem of dense edge areas can be achieved by windowing. The proposed approach consists in defining a window $W(\varepsilon_{eq})$ centered on each image point for which the potential is to be computed, and defining the area within which surrounding edge points have to be considered.

The size of the window must trade-off between two aspects. On one side, since the windowing procedure results in an approximation of the field, it is desirable that the achieved function is similar to the actual one. On the other side, the necessity of reducing the impact of locally dense edge areas on the potential function requires a limitation of the radius of influence relevant to a charged element (a sort of interdiction zone).



(a)

(b)

Figure 1: *BEPF obtained from a binary image: (a) original image; (b) edge potential function for $\varepsilon_{eq} = 0.05$.*

As to the first point, it is important to notice that the effect of a charged element on the potential of a given point is inversely proportional to the distance of the charge itself. Therefore, depending on the value of ε_{eq} there is a distance value after which the effect of the charge on the potential is negligible. If we assume this distance to be proportional to the window size, the effect of windowing is quite negligible except for the computational gain. As to the second point, the effect of the interdiction zone is inversely proportional to the window size, so that the use of a small window can significantly reduce the problem of dense edge areas. Taking into account the two problems, and considering that typical values for the equivalent permittivity are in the range $[0.01 \div 0.2]$, reasonable values for the window size are in the range $[6 \div 15]$ pixels. The above permittivity values were heuristically defined taking into consideration the effect of ε_{eq} on the range of attractivity of the contour.

Fig. 2 shows the difference between the potential function computed with Eq. 3 and the approximated one: it is to be observed that the windowed EPF is characterized by a larger response near to the contour points and a steeper transition region.

3. THE GENETIC ALGORITHM

We have now to define how the EPF can be used to perform the optimal match between test image (the target) and model image (the sketch). Since in general we accept that the sketched object is present in the image at any position, rotation and scaling, it is necessary to define a set of transformations of the original contour and a metric quantifying the matching. To this end, the following geometrical operators are taken into account:

- the rotation, ϑ;
- the translation (along the horizontal, t_x, and vertical, t_y, directions, respectively);
- the scaling, (along the horizontal, t_w, and vertical, t_h, directions, respectively);

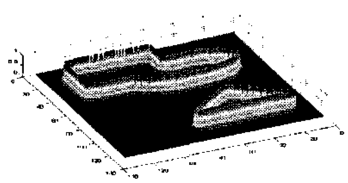


Figure 2. Result of approximated computation of BEPF using windowing, applied to the same image of Fig. 1.a (ε_{eq} = 0.05, W=8)

By considering these operators, the original sketch (c_0) is successively modified obtaining different instances of the contour (c_k, k = 1, ..., K), which must be fitted within the potential field. As far as the definition of a suitable metric is concerned, the following "fitting" measure is defined by considering the EPF:

$$f(c)=\frac{1}{N^{(c)}}\sum_{n^{(c)}=1}^{N^{(c)}}\left\{EPF(x_n^{(c)},y_n^{(c)})\right\} \tag{4}$$

where $n^{(c)}$ is the n-th pixel of the c-th contour, and $N^{(c)}$ is the total number of pixels of c-th contour. Accordingly, the optimal matching (corresponding to the set of transformations that provides the maximum potential along the contour c_{opt}, among all possible transformations) is obtained when the *fitness* function is maximized.

Generally speaking, the fitness measure in Eq. 4 is an example of a multi-modal and nonlinear function. Consequently, a suitable meta-heuristic (providing the rule of change of the original sketch and minimizing Eq. 5), able to avoid the solution be trapped in local minima, has to be chosen.

Genetic Algorithms [6] are searching processes modeled on the concepts of natural selection and genetics. Their basic principles were first introduced by Holland in 1975 [7] and extended to functional optimization by De Jong [8]. GAs have been recently employed with success in a variety of engineering applications. What makes GA's immediately attractive is that they can be easily applied to problems involving non-differentiable functions and discrete as well as continuous spaces. These are qualities shared by other techniques, but GAs also exhibits an intrinsic parallelism.

The key items in designing a GA-based inversion procedure are:

- the representation of the solution, $c \Rightarrow \tilde{c}$;
- the design of evolutionary operators responsible for the generation of the trial solution succession, $\{c_k, k = 1, \dots, K\}$;
- the evolutionary procedure.

Different choices result in different procedures able to efficiently deal with different problems. As far as the shaped-based retrieval of images is concerned, a customized binary version of the algorithm has been taken into account. In this case, the algorithm operates on a coding of the problem parameters. The GA, when applied to minimize the functional $f(c)$ (see Eq. 5), requires the definition of a population of trial solutions

$$P_0=\left\{c_0^{(p)};p=1,\dots,P\right\} \tag{5}$$

being P the dimension of the trial solution population. Iteratively (being k the iteration number), the solutions are ranked according their *fitness* measures

$$F_k=\left\{f\left(c_k^{(p)}\right)p=1,\dots,P\right\};\ k=0,\dots,K \tag{6}$$

and, following the classical binary-coded version of the GA [6], coded in strings of $N=\sum_{i=1}^{N_{par}}\{\log_2[Q_i]\}$ bits, being Q_i the number of quantization levels used for i-*th* of the N_{par} = 5 unknowns, $\{\vartheta, t_x, t_y, t_w, t_h\}$,

$$\Gamma_k=\left\{\tilde{c}_k^{(p)};p=1,\dots,P\right\} \qquad k=0,\dots,K \tag{7}$$

At this point, new populations of trial solutions are iteratively obtained by applying the genetic operators (selection, crossover and mutation) according to a steady-state strategy [9]. At each iteration, a *mating pool* is chosen by means of a tournament selection procedure

$$\Gamma_{k(\delta)}=\delta\{\Gamma_k\} \tag{8}$$

where δ indicates the selection operator. Then, a new population is generated applying, in a probabilistic way, the two-point binary crossover, ξ, and the standard binary mutation, ζ, [10]

$$\begin{gathered}\Gamma_{k+1}=\Gamma_{k(\xi)}\cup\Gamma_{k(\varsigma)} \\ \Gamma_{k(\xi)}=\xi\left\{\Gamma_{k(S)}\right\} \qquad \Gamma_{k(\varsigma)}=\varsigma\left\{\Gamma_{k(S)}\right\}\end{gathered} \tag{9}$$

The genetic operators are iteratively applied corresponding to their probabilities. Crossover is applied with a probability P_c, mutation is carried out with a probability P_m, the rest of the individuals are reproduced. The iterative generation process stops when the stationary condition is fulfilled

$$\frac{\sum_{i=1}^{I}\left|f_k^* - f_{k-i}^*\right|}{I} < p_{conv} \qquad (10)$$

where , $f_k^* = \max_{\substack{p=1,\dots,P \\ h=1,\dots,k}} \left\{ f\left(c_h^{(p)}\right) \right\}$, I is a fixed number of iterations, and $P_{conv} \in [0,1]$ is the convergence threshold.

4. EXPERIMENTAL RESULTS

The proposed technique has been widely tested to validate its performance. In this section, two examples are reported, in order to allow a better evaluation of the capabilities of the method. The first example (Fig. 3) shows the effectiveness of the matching procedure in detecting the presence of a user-given sketch within a binary image representing some natural object shapes (3.a). Two kinds of sketches are used in the query: the former (3.b) is directly extracted from the image, while the latter (3.c) is hand drawn by the user, thus with a shape not perfectly overlapped with the target. Windowed EPFs were adopted in both cases, with a window size equal to 10. The results of the matching procedure for the two sketches are provided in Figs. 3.d-e, respectively.

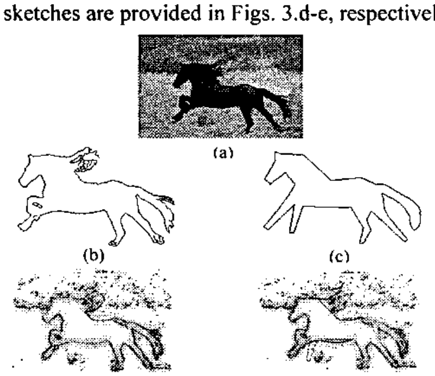


Figure 3. Application of the proposed method to a binary test image. (a) original; (b) sketch extracted from the original; (c) hand-drawn sketch; (d-e) matching achieved with sketches (b) and (c), respectively.

The second example (Fig. 4) reports the result achieved in a shape-based ranking procedure applied to a large image database (around 500 natural and synthetic images with various subjects). Here, the maximum fitness value (MFV) is used as a confidence parameter to sort the images on the basis of the possible presence of the sketched object.

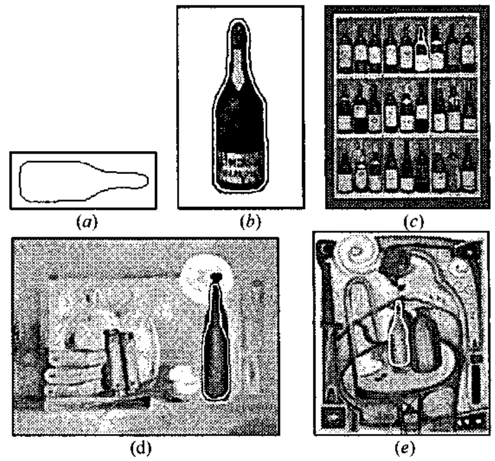


Figure 4.(a) sketch (b, c, d, e) first four matching images in the DB, corresponding to MFV values of 0.89, 0.57, 0.56, 0.54, respectively. In each image the sketch is superimposed at the position, rotation and scaling that provided the maximum fitness.

5. REFERENCES

[1] Arnold W.M. Smeulders, Marcel Worring, Simone Santini, Amarnath Gupta, Ramesh Jain: "Content-based image Retrieval at the End of the Early Years". IEEE Trans. on PAMI, Vol. 22, No.12, Dec., 2000.

[2] A. Del Bimbo, and P. Pala, "Visual Image Retrieval by Elastic Matching of User Sketches," IEEE Trans. on PAMI, Vol. 19, No. 2, Feb. 1997.

[3] P.W.M. Tsang: "A genetic Algorithm for Affine Invariant Object Shape Recognition", Proc. 1st Int. Conf. Genetic Algorithms in Engineering Systems: Innovations and Applications, 1995.

[4] T. Kawaguchi, R. Nagata, T. Simozaki: "Detection of Target Models in 2D Images by Line-Based Matching and a Genetic Algorithm", Proc. IEEE Int. Conf. on Image Processing, 1999.

[5] T. Kawaguchi, M. Nagao: "Recognition of Occluded Objects by a Genetic Algorithm", Proc. 14th IEEE Int. Conf. on Pattern Recognition, 1998.

[6] D.E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, Reading, MA, 1989

[7] J. H. Holland, *Adaptation in Natural and Artificial Systems*, Univ. Michigan Press, Ann Arbor, 1975

[8] K. A. De Jong, An analysis of the behavior of a class of genetic adaptive systems, Ph.D. Dissertation, Univ. Michigan, 1975

[9] J. M. Johnson and Y. Rahmat-Samii, "Genetic Algorithms in engineering electromagnetics," IEEE Trans. on Antennas and Progat. Magazine, Vol. 39, No. 4, August 1997.

[10] M. Pastorino, A. Massa and S. Caorsi, "A microwave inverse scattering technique for image reconstruction based on a genetic algorithm", IEEE Trans. on Instrumentation and Measurement, vol. 49, no. 3, June 2000.