

OBJECT SIMPLIFICATION USING A SKELETON-BASED WEIGHT FUNCTION

András Hajdu, Charalambos Giamas, Ioannis Pitas

Artificial Intelligence and Information Analysis Laboratory,
Department of Informatics, Aristotle University of Thessaloniki, Greece
pitas@aia.csd.auth.gr

ABSTRACT

In this paper, we present a novel method for template simplification, where the template is used to find interesting objects within an image. In this way, we improve computational performance since less template points are matched using a simplified template. Moreover, we increase the reliability of the matching as we keep template points with focusing on the main shape behavior (skeleton) of the template. The theoretical background of the simplification is derived through the centroidal Voronoi tessellation framework. The efficiency of the proposed approach is demonstrated with detecting human appearance in thermal images.

1. INTRODUCTION

Object detection and classification are challenging problems in digital image analysis and computer vision. Usually, some image features are extracted, and a matching procedure is applied afterwards to find occurrences of pre-defined templates. In 2D, these templates can be primarily classified as contour or region-like objects. The matching of the templates can be performed either directly (with considering all the pixels of the template), or through the extraction of some shape descriptors. In the case of direct pixel matching, it is a natural motivation to gain faster computation via the simplification of the description of the objects to be matched, by representing them with less number of points.

In this paper, we propose a novel way to perform the simplification of the object description to gain faster template matching. Our approach, which is based on centroidal Voronoi tessellation (CVT) framework [1, 2], can be applied to any dimensions. With this work, we continue some former investigations [3] focusing on contour-like templates. Using a similar theoretical approach as in [3], we explain how we can take advantage of CVT-based algorithms in representing region-like objects with less number of points dedicated to object matching purposes. Our main contribution here is to define and apply weight functions that leads to such a simplification, which preserves the main shape behavior of the template represented by its skeleton. In this way, we can allow larger difference between the boundary of the target object and the template with still having a reliable matching.

The structure of the paper is as follows. In section 2 we recall the necessary elements of the CVT framework. We also explain our object matching method which is based on a direct check through all the points of the template. Then, section 3 describes how the objects to be matched can be simplified using CVT algorithms accumulating the existence of a shape-based weight function for point selection. Section 4 contains our experimental results regarding the proposed simplification in a human silhouette matching scenario. We end up with conclusions in section 5.

2. CENTROIDAL VORONOI TESSELLATIONS

The centroidal Voronoi tessellation framework [1] provides the possibility to select a given number of points of a set which are distributed according to some weight function (generally the uniform one). Now we give a brief theoretic and algorithmic recall of the CVT framework to explain how a specific weight function can be considered to perform object simplification accordingly. The simplification problem can be formulated in the following way:

Problem: Let $A \subseteq B \subseteq \mathbb{R}^m$, such that A is compact and convex and B is bounded. Moreover, let $K \in \mathbb{N}$, and $\varrho : B \rightarrow \mathbb{R}_{\geq 0}$. Find the set of points $A' = \{\mathbf{y}_i \in A \mid i = 1, \dots, K\}$, which minimizes

$$\sum_{i=1}^K \int_{V_i(A')} \varrho(\mathbf{y}) |\mathbf{y} - \mathbf{y}_i|^2 d\mathbf{y}, \quad (1)$$

where $\{V_i(A')\}_{i=1}^K$ denote the Voronoi tessellation of B generated by A' .

We note here that problem (1) is well investigated when $A = B$. It has been shown that the optimal solution should define a centroidal Voronoi tessellation on A [1], which means that the generators of the Voronoi cells are also the mass centers (centroids) of them, respectively. This kind of distribution of points can be achieved using CVT algorithms [1, 4, 5, 6].

Research was supported by the project SHARE: Mobile Support for Rescue Forces, Integrating Multiple Modes of Interaction, EU FP6 Information Society Technologies, Contract Number FP6-004218.

The CVT approach together with the generating algorithms were extended to the tessellation of surfaces by considering constrained CVT [2]. Later, in [3], the CVT framework was generalized to simplify sets also with a consideration to the influence of a container region, that is, the $A \subset B$ case. This approach is called region-based CVT (RCVT), and the main intention here was to simplify objects to achieve faster template matching with a less number of template points. The main scope of [3] was to simplify contour-like objects, though the theory was given more generally, also for region-like objects. As it was shown, the general theory remains valid also for non-uniform weight functions, however, their use was not considered. To compute RCVT we may consider e.g. iterative statistical methods based on random sampling [4, 7]. They are initialized with a random selection of K points.

Random sampling algorithm for RCVT (see also [3])

Step 1: Choose a $q \in \mathbb{N}$ and constants $\alpha_1, \alpha_2, \beta_1, \beta_2$, such that $\alpha_2, \beta_2 > 0$, $\alpha_1 + \alpha_2 = \beta_1 + \beta_2 = 1$; choose an initial set of K points $\mathbf{z}_1, \dots, \mathbf{z}_K$ in A , e.g. by using a Monte Carlo method; set $j_i = 1$ for $i = 1, \dots, K$;

Step 2: Choose q points $\mathbf{y}_1, \dots, \mathbf{y}_q$ in B at random, e.g. by a Monte Carlo method, according to some probability density function (in this description uniformly);

Step 3: For $i = 1, \dots, K$, gather together in the set W_i all sampling points \mathbf{y}_r closest to \mathbf{z}_i among $\mathbf{z}_1, \dots, \mathbf{z}_K$, i.e. lying in the Voronoi region of \mathbf{z}_i w.r.t. B ; if the set W_i is empty, do nothing; otherwise, compute the \mathbf{u}_i average of the set W_i and set

$$\mathbf{z}_i \leftarrow \frac{(\alpha_1 j_i + \beta_1) \mathbf{z}_i + (\alpha_2 j_i + \beta_2) \mathbf{u}_i}{j_i + 1}, \quad j_i \leftarrow j_i + 1. \quad (2)$$

The new set of $\mathbf{z}_1, \dots, \mathbf{z}_K$ along with the unchanged \mathbf{z}_j 's (i.e. when W_i is empty), form the new set of points $\mathbf{z}_1, \dots, \mathbf{z}_K$;

Step 4: If for some i , $\mathbf{z}_i \notin A$ then $\mathbf{z}_i \leftarrow \mathbf{z}$, where $d(\mathbf{z}_i, \mathbf{z}) = \min_{\mathbf{y} \in A} \{d(\mathbf{z}_i, \mathbf{y})\}$, i.e. \mathbf{z} is the nearest point to \mathbf{z}_i in A .

Step 5: If the new points meet some convergence criterion, terminate; otherwise, return to step 1.

As it is shown in Figure 1, as the a naive approach, we can simplify the template uniformly. Here we can follow two ways: we select object points according to a uniformity rule with respect to pixel coordinates (e.g. both the coordinates are equal to 0 modulo 4), or we consider a uniform weight function ϱ in the RCVT framework. We note that these representations are pure simplifications with completely reflecting the original shape behavior.

Our motivation is to introduce weight functions for which the CVT/RCVT object simplification lead to natural improvements in the case of region-like object matching.

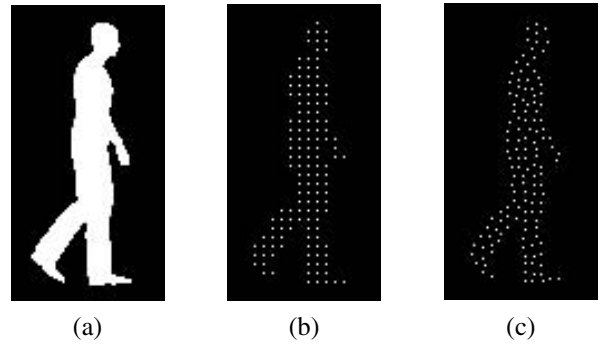


Fig. 1. The result of the RCVT algorithm in 2D; (a) the object A to be simplified, (b) trivial uniform sampling according to pixel coordinates, (c) uniform sampling based on RCVT.

3. SKELETON-BASED SIMPLIFICATION

To realize a more adaptive approach than the naive ones discussed above, it is also possible to focus more on some object regions during its simplification, e.g. if we expect larger uncertainty closer to the boundary of the object in the matching procedure. In this case, we can define a weight function $\tilde{\varrho}$ according to problem (1), which concentrates on the object morphological skeleton [8]. Thus, for every $\mathbf{x} \in B$, let us define:

$$\tilde{\varrho}(\mathbf{x}) = \frac{d(\mathbf{x}, \overline{B^c})}{d(\mathbf{x}, \text{sk}(A)) + d(\mathbf{x}, \overline{B^c})}, \quad (3)$$

where $\text{sk}(A)$ is the skeleton of A , B^c is the complement of B , and d is some common distance function. Note that for all $\mathbf{x} \in B$, $0 \leq \tilde{\varrho}(\mathbf{x}) \leq 1$, such that $\tilde{\varrho}$ vanishes at the boundary of B , then monotonously increases till reaching $\text{sk}(A)$, where it takes value 1. In this way, the weight for the points within B is adjusted according to their relative distance from the "center" of A and from $\overline{B^c}$. Now, to derive a weight (density) function over B , let:

$$\varrho(\mathbf{x}) = \frac{\tilde{\varrho}(\mathbf{x})}{\int_B \tilde{\varrho}(\mathbf{y}) d\mathbf{y}}, \quad \text{for } \mathbf{x} \in B. \quad (4)$$

In applications, this weight function can be easily derived using the distance maps [9] of $\text{sk}(A)$, and B^c to approximate d . Figure 2 depicts such an example, where $A = B$, the basic CVT case is used. We also found the pruning (removal of small branches) of the skeleton to be useful.

4. EXPERIMENTAL RESULTS

To be able to begin with the template matching, first we need to extract the target object from the input image. In our currently developed system, one of the desired tasks is the detection of human (victim or firefighter) appearance within thermal videos from a rescue scenario. To reach this aim, we considered a fuzzy segmentation technique [10], which is known

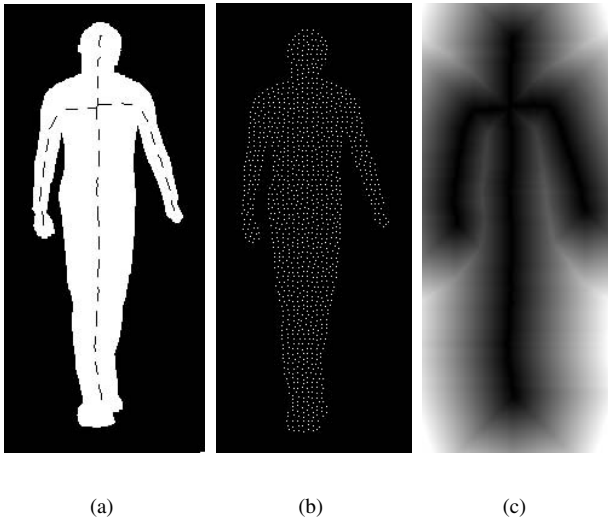


Fig. 2. Simplification result of an object according to a weight function concentrating on its skeleton; (a) target object $A = B$ ($sk(A)$ is shown by dashed line), (b) result of simplification in case of uniform weighting, (c) distance map of $sk(A)$, (d) distance map of $\overline{B^c}$, (e) weight function $\tilde{\varrho}$ (higher intensities show larger weight values), (f) result of CVT simplification using ϱ derived from $\tilde{\varrho}$.

to be robust also for medical (e.g. CT) images. The extracted binary region usually needs some simple post-processing to smoothen the boundary, eliminate gaps and holes, etc. These minor refinements can be achieved by some classic elements of mathematical morphology [8]. Some input test images together with the result of the fuzzy segmentation is shown in Figure 3.

To create templates to be matched, we created simulated

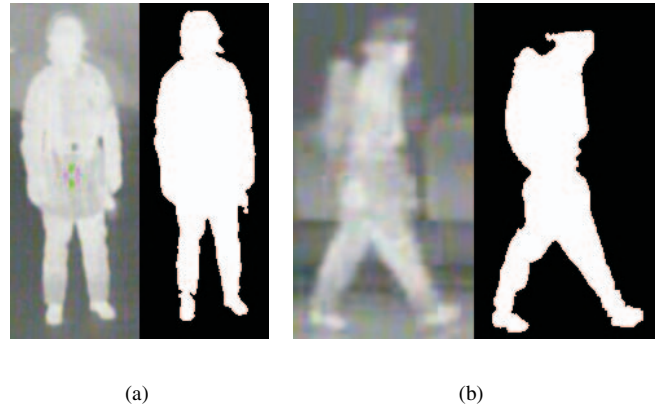


Fig. 3. The result of the fuzzy segmentation; (a) standing human pose, (b) walking human pose.

(artificial) data using the realistic 3D human motion software Poser [®] [11]. As it is not a crucial point in our present experiments, we ignored all the geometric transformation issues on how to align the target and template objects. As a guideline, we recommend to consult with [12] to gather information on a robust search of the affine parameter space.

To demonstrate the main idea of our approach, we show the performance of such templates that do not completely fit the target. With the skeleton-based simplification, in the case of a regular human pose/motion (like standing/walking), we can expect better matching for more template elements, since the skeleton does not change drastically with the boundary e.g. for close phases of the same motion. The goodness of fit value of a template against the target object is calculated as the percentage of the matching pixels of the template at the best matching position. Figure 4 and 5 show the best matching positions of a standing and walking template together with their discussed simplifications, respectively. The skeleton of the original template is also marked.

We can see that the skeleton-based simplification has better performance than the uniform ones, since the main skeleton of the target objects and the templates did not differ that much. The proper quantitative results are given in Table 1.

	Simplification			
	No (original)	Unifom (trivial)	Unifom (CVT)	Weighted (skeleton)
Standing	80,9%	81,2%	81,1%	94,0%
Walking	85,8%	86,2%	86,2%	94,6%

Table 1. Goodness of fit of simplified templates given in terms of the percentage of matching template points.

This analysis also validates the natural assumption that the uniformly sampled template performs similarly to the original

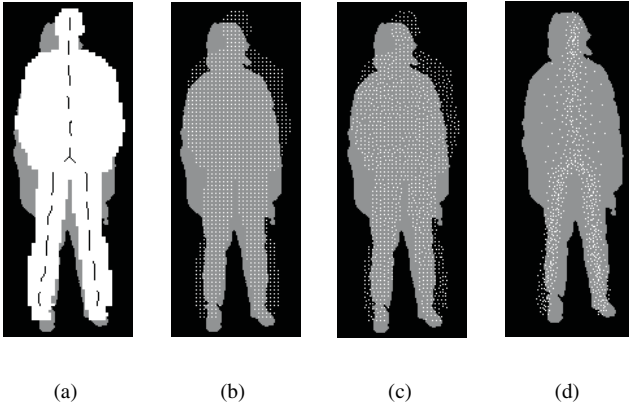


Fig. 4. Best matching position (shown in white) for standing template using (a) original template (skeleton is marked), (b) trivial uniform simplification, (c) CVT-based uniform simplification, (d) skeleton-based simplification. Target object is shown in gray.

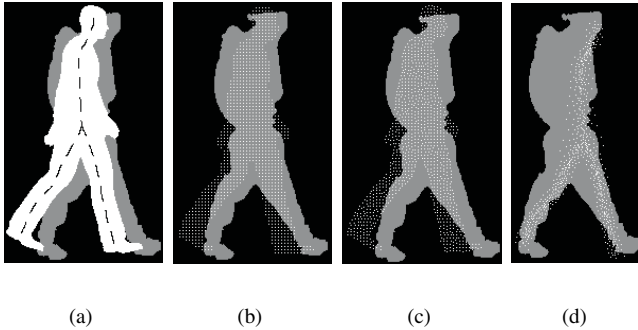


Fig. 5. Best matching position for walking template using (a) original template, (b) trivial uniform simplification, (c) CVT-based uniform simplification, (d) skeleton-based simplification.

one.

5. CONCLUSION

In this paper, we proposed a novel simplification approach of objects designated to region-based object matching purposes. In this way, we gain computational improvement, since less number of points are involved in the matching process. The simplification can be done using centroidal Voronoi tessellation framework with a weight function concentrating on the skeleton of the object. We can expect to capture some regular motions with using less number of templates with the computation time is also reduced by this simplification. To

demonstrate this feature, we presented some experimental results. The approach can be easily extended to higher dimension, and to weight functions for other tasks.

6. REFERENCES

- [1] Q. Du, V. Faber, and M. Gunzburger, Centroidal Voronoi tessellations: Applications and algorithms, *SIAM Review*, 41:637-676, 1999.
- [2] Q. Du, M. Gunzburger, and L. Ju, Constrained Centroidal Voronoi tessellations on general surfaces, *SIAM J. Scientific Comp.*, 24:1488-1506, 2003.
- [3] A. Hajdu, A. Roubies, and I. Pitas, Optimized chamfer matching for snake-based image contour representations, *ICME 2006*, Toronto, Canada, 2006.
- [4] L. Ju, Q. Du and M. Gunzburger, Probabilistic methods for centroidal Voronoi tessellations and their parallel implementations, *Parallel Computing* 28, 1477-1500, 2002.
- [5] S. Lloyd, Least square quantization in PCM, *IEEE Trans. on Information Theory*, 28:129-137, 1982.
- [6] J. MacQueen, Some methods for classification and analysis of multivariate observations, *Fifth Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley, I:281-297, 1967.
- [7] Q. Du, M. Emelianenko, and L. Ju, Convergence properties of the Lloyd algorithm for computing the Centroidal Voronoi Tessellations, *SIAM J. Num. An.*, 44(1):102-119, 2006.
- [8] J. Serra, *Image Analysis and Mathematical Morphology*, Academic Press, New York, 1988.
- [9] G. Borgefors, Hierarchical chamfer matching: A parametric edge matching algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(6):849-865, 1988.
- [10] L.G. Nyúl, A.X. Falcao, and J.K. Udupa, Fuzzy-connected 3D image segmentation at interactive speeds, *Graphical Models*, 64(5):259-281, 2002.
- [11] <http://www.e-frontier.com/>
- [12] W. Rucklidge, Locating objects using the Hausdorff distance, *International Conference on Computer Vision*, pp. 457-464, 1995.