ABSTRACT

Residual operators analyze the evolution of an image subject to the application of a series of transformations, for example a series of openings of increasing size. When a significant object is filtered out by a transformation corresponding to its size, an important residue is observed. Maximal residues are kept for each pixel, indicating the most significant objects present in the image. Different families of operators have been used in the literature: morphological openings or closings, attribute openings or openings by reconstruction. In this paper we propose to compute residues on a hierarchy of partitions, computing the differences between regions at different hierarchical levels based on the classical earth mover’s distance. The advantage of our approach is that it is autodual and generic as it can be applied with any hierarchical approach.

Index Terms— Numerical residues, ultimate opening, waterfall, P algorithm, image segmentation, mathematical morphology, earth’s mover distance.

1. INTRODUCTION

A residue is defined as the difference between two operators. Among residual operators we find the morphological gradient, the top-hat transformation, thinnings or the skeleton transform, all of them extensively used in the literature. More recently Serge Beucher generalized the residue of a family of operators to gray level images introducing the quasi-distance and the ultimate opening operators [1]. Previous versions of ultimate opening were independently proposed in the literature under the name of adaptive opening by Vogt [2] or morphological profiles by Pesaresi et al. [3, 4]. Successful applications based on these operators have been developed such as rock analysis [5], automatic text localization [6], façade segmentation [7] and remote sensing approaches [8, 9, 10].

In practice, residual operators based on morphological openings (or closings) or attribute openings with different attributes have been used. The choice of the attribute has a strong influence in the final result. Moreover those operators deal with openings or closings, focusing on either bright or dark objects. Some authors propose to combine both polarities choosing for each pixel the polarity leading to the highest contrast estimation [4, 11]. Thus bright and dark objects are correctly processed but intermediate gray level regions can be missed. A similar idea is used by Maximally Stable Extreme Regions (MSER) [12]. The image is quantized with a given threshold step. A min-tree of the resulting image is built. The area difference between connected components of consecutive thresholds is computed. Significant regions are chosen when this area difference constitutes a regional minimum. The choice of the threshold step decides which are the regions to be compared and is a critical parameter of the algorithm.

In this paper we propose a residual approach based on a hierarchy, dealing with both polarities simultaneously and correctly considering intermediate gray level regions.

2. HIERARCHY

Let $H = \{P_0, P_1, ... P_N\}$ be a hierarchy of nested partitions of image $I$. $P_0$ corresponds to the finest partition: e.g the set of pixels of the input image, the watershed result (usually over-segmented) or any other input partition. $P_N$ corresponds to a partition with a single region for the whole image. $P_i$ is a simplification of $P_{i-1}$ which means that a region $R_{ji}$ of $P_i$ is the union of a set of regions $R_{ki}$ of $P_{i-1}$ such that $R_{ki} \subseteq R_{ji}$:

$$R_{ji} = \{ \cup_{R_{ki}} \mid \forall k \; R_{ki} \subseteq R_{ji} \}$$

$R_{ji}$ is said to be the parent of regions $R_{ki}$ contained in it.

The proposed approach is general: any hierarchical approach can be used. In order to illustrate it we use watershed-based hierarchies. The Watershed algorithm usually leads to an over-segmentation of the image. Several hierarchical approaches have been proposed to overcome this problem. In this paper we focus on waterfall hierarchy [13] and its enhancement, the P algorithm [14]. The waterfall [13] is a watershed-based hierarchical segmentation approach. It consists in two steps:

- first, each region is filled with the value of the lowest pass point of its frontier. The pass point is the gray level
value of the pixel where, during the flooding process associated with the watershed, neighbouring “lakes” (regions) meet for the first time. A morphological reconstruction may be used for this purpose.

- second, the watershed of the resulting image is computed.

In the example of Figure 1 the watershed lines are indicated by arrows and only solid line arrows will be preserved by the waterfall. The process may be iterated until a single region covers the whole image, establishing a hierarchy among the frontiers produced by the watershed. An efficient graph-based waterfall algorithm is presented in [15].

In figure 1(b), dotted arrows are boundaries removed by the waterfall and blue dotted arrows are those re-introduced by P algorithm.

3. RESIDUAL APPROACH ON A HIERARCHY

Ultimate opening identifies important regions thanks to the strong gray level change they produce when they are filtered out by an opening of its corresponding size. In order to introduce the residual operator linked to a hierarchy, we have to estimate the distance between child and parent nodes within a hierarchy of nested partitions. The mean grey level difference is not a good choice because it is meaningless for the highest levels of the hierarchy. We propose to use a histogram distance computed between a region, \( R_{i-1}^k \), and its parent, \( R_i^j \):

\[
Res(R_{i-1}^k) = \text{distance}(R_{i-1}^k, R_i^j)
\]

The Earth Mover’s Distance (EMD) between two distributions is defined as the minimal cost that must be paid to transform one distribution into another. It is supposed to match perceptual similarity better than other distribution distances used for image retrieval [16]. The EMD avoids quantization and other binning problems typical of histograms and allows partial matching. When used to compare distributions with the same overall mass, the EMD is a true metric.

We use EMD in order to estimate the distance between regions of consecutive hierarchical levels. If the fusion is performed within a homogeneous region the distance \( \text{distance}(R_{i-1}^k, R_i^j) \) is low whereas this distance will be high if distinctive regions are merged. The idea is to detect significant changes in the hierarchy and avoid them, keeping the regions before critical mergings. Then, the hierarchical level is chosen locally, as residues are computed for each region of \( P_{i-1} \). The decision is taken region by region. Thus the final partition will be made up combining different hierarchical levels.

The principle is illustrated in figure 2. In this example, the finest partition has 6 regions (A-F) and the hierarchy has 3 levels. The hierarchy is represented as a tree. The process starts from the finest partition. Residual image is initialized with the distance between leave regions and their parents. Then higher hierarchical levels are considered. The relevance of a region is estimated as the distance with its parent. Then children’s and parent’s relevances are compared:

- If children are more important than their parent, children are kept (regions A, B and C in our example).
- If the parent is more relevant than its children, the parent is chosen (regions are merged). This is the case for region H, including regions D, E and F.

The process is iterated for higher hierarchical levels. In figure 2(b) plain nodes are kept while dashed ones are removed. Figure 2(c) illustrates the final result. Each time that a region

![waterfall principle](image1)

![P algorithm principle](image2)
$R_k^i$ is kept, its corresponding residue $Res(R_k^i)$ is stored in the residual image ($T(p)$). The residual image $T(p)$ for a pixel $p$ is then defined as:

$$T(p) = \max\{Res(R_k^i) \mid p \subset R_k^{i-1} \land i \in \{1, \ldots, N\}\}$$

(1)

$T(p)$ conveys contrast information of a region with its neighborhood.

In order to get a correct segmentation result, it is important to have significant regions at a given level of the hierarchy. It is known that waterfall produces few hierarchical levels, with strong simplifications between consecutive levels. $P$ algorithm produces more hierarchical levels than the waterfall hierarchy, more meaningful partitions and a non-null partition at its highest level. In order to complete the hierarchy, the two last waterfall levels are added. Figure 4 shows successive $P$ algorithm hierarchical levels. Figures 4(h) and (i) show the final partition and the corresponding associated contrast. We can observe in $P_3$ the correct segmentation of the square area mentioned above, and correctly selected in the final partition 4(h).

More examples comparing $P$ algorithm, waterfall residual approach and $P$ algorithm residual approach are provided in figure 5. We can observe that $P$ algorithm residual approach outperforms $P$ algorithm itself as well as waterfall-based residual approach.

5. CONCLUSION AND PERSPECTIVES

In this paper we generalize the residual approach to a segmentation hierarchy. We demonstrate that the residual approach is able to identify significant regions from the hierarchical segmentation approach. The number of regions adapts
automatically to the image complexity. Moreover, additional contrast information is provided. This contrast is based on the earth mover’s distance. We show that P algorithm residual approach outperforms P algorithm itself as well as waterfall residual approach. In the future fast implementations of the proposed algorithm will be addressed and color distances will be considered.
6. REFERENCES


[14] Serge Beucher and Beatriz Marcotegui, “P algorithm, a dramatic enhancement of the waterfall transformation.,” http://hal-ensmp.archives-ouvertes.fr/hal-00835019/PDF/P-Algorithm_SB_BM.pdf, 2009-09.

