MORPHOLOGICAL RESIDUAL REPRESENTATIONS OF SIGNALS

Dinu Coltuc[1], Ioannis Pitas[2]

[1]Research Institute for Electrical Engineering, Bucuresti, Romania

[2]Department of Informatics, University of Thessaloniki, Thessaloniki 540 06, P.O. Box 451, Greece

Abstract. A general approach for residual representation of 1D and 2D signals is defined. Signals are reconstructed as a sum of components that are recursively determined by using a constructive transform. Several morphological constructive transforms are proposed and their corresponding representations are discussed. The use of residual representations in signal and image compression is investigated with promising results.

1 Introduction

A problem of considerable concern in signal/image processing and analysis is the representation of the data in terms of meaningful geometrical patterns that are characterized by shape and size. The shape-size relation is reflected by the scale concept. The nonlinear approach (based on mathematical morphology) for multiscale representation is strongly advocated in [1]; scale is defined as the smallest size of a shape pattern (generated by a prototype pattern of unit size) that can "fit" inside the signal. This paper proposes a general framework for 1D and 2D signal representation in series of simple components that are recursively extracted by a constructive transform. The representation scheme is meant to be very general. Scale information is embedded in the representation by an appropriate selection of the constructive transform, i.e., the multiscale opening. Morphological shape decomposition (MSD) [2] is a particular case of our representation scheme. Refinements of the constructive transform and adaptive schemes are further discussed.

2 Residual representation

Let f be an 1D or 2D signal, $f \geq 0$ on its domain. Our aim is to represent f as a set of components, $\{f_i, i = 1, 2, \ldots, \}$, such that $f = \sum_i f_i$. The components are extracted from a family of residues of f by means of a transform, called constructive transform. The residue of two transforms, ϕ and ψ, is defined as $R_{\{\phi,\psi\}}(f) = \phi(f) - \psi(f)$. The general representation scheme is provided by the following definitions:

Definition 1. *A transform γ is a constructive transform if it is antiextensive and positively defined.*

Definition 2. *The family of residues of f with respect to γ is $\{r_i, i = 1, 2, \ldots, \}$, where $r_1 = f$ and $r_{i+1} = R_{\{r_i, \gamma(r_i)\}}$ until $\gamma(r_i)$ is a null signal.*

Definition 3. *The set $\{f_i\}$, $f_i = \gamma(r_i)$ is the residual representation of f with respect to the constructive transform γ. The reconstruction of f, denoted by f^*, is the summation of the components of the representation set, $f^* = \sum_i f_i$. The partial j-components reconstruction is given by $f_j^* = \sum_{i=1}^{j} f_i$.*

The family of residues is, from a morphological point of view, a family of residues of two transforms: the identity transform $I(r) = r$ and the constructive transform γ. The residue family is recursively generated by a procedure uniquely determined by the constructive transform. The properties of γ assures a representation by a series of nonincreasing components, $f_{i+1} = \gamma(r_{i+1}) \leq \gamma(r_i) = f_i$, where the order relation on functions is $f \leq g$ if $f(x) \leq g(x)$ is valid everywhere. The representation of f is unique, provided that γ is a well-defined transform. This property is fundamental for any representation scheme.

The representation of f in arbitrary levels of detail by truncating the representation sequence to a predefined number j of components is an useful property. The reconstruction is nondecreasing by definition. Since $f = f_j^* + r_{j+1}$, the reconstructed f_j^* equals the original when the final residue r_{j+1} is empty. If r_{j+1} is not empty and $\gamma(r_j)$ is a null function, the reconstruction is not converging and its error has to be determined (e.g., by L_2 norm).

A large number of transforms fulfil the above-mentioned constraints. Thus, many residual representations exist. The problem is to determine constructive transforms that generate representations of interest for signal processing and analysis, i.e., possessing other properties as well, e.g., compactness, translation, scale and rotation invariance.

3 Morphological representations

Morphological erosions and openings are antiextensive transforms if structuring functions satisfy $g(0) > 0$. Furthermore, they are constructive transforms if the region of support of the erosion $[f \ominus kg^s](x)$ is limited to those points x of the domain for which $[f \ominus kg^s](x) \geq 0$. Let r_{kg} be the family of openings of r with kg, the homothetics of a structuring function g. As k increases, the openings are more simplified versions of the original function and the details that are less than kg are discarded. This reason suggests us the selection of the maximal scale opening as the constructive transform. If κ is the corresponding maximal value of k, the constructive transform is $\gamma(r) = \max_k(r_{kg}) = r_{\kappa g}$. Following the approach of section 2, the definitions of binary or graylevel morphological shape decomposition are recovered [2]-[4]. The family of residues is $r_{i+1} = R_{\{r_i, r_{\kappa_i g}\}}$ and the components of the representation are:

$$f_i = \gamma(r_i) = (r \ominus \kappa_i g^s) \oplus \kappa_i g. \tag{1}$$

Information on the shape of the signal is captured by g while κ_i provides information on size. The components are maximal functions in the sense that they describe the largest scaled version of the structuring function g that can still be fitted inside the current residue. The maximal structuring function is an invariant for the residue. The support of the components f_i are the functions l_i, $l_i = r_i \ominus \kappa_i g^s$. By using l_i functions, one can get $f_i = l_i \oplus \kappa_i g$. The function l_i represents the loci of the centers of the maximal structuring function in the current residue, l_i is positive on its domain. It is easy to see that the support of each component is the subset of the skeleton of the residue having the maximum quench function. The reconstructed f is given by:

$$f^* = \sum_i f_i = \sum_i (r_i \ominus \kappa_i g^s) \oplus \kappa_i g = \sum_i l_i \oplus \kappa_i g \tag{2}$$

The reconstruction of a test signal by using the multiscale opening constructive transform is shown in Fig. 1. The quality of the final reconstruction is high, its peak-to-peak signal to noise ratio (PSNR) is 36.88 dB.

The morphological representation is translation invariant and, in the continuous setting, it is scale invariant and rotation invariant (if the structuring function is rotation invariant). The quality of the reconstruction depends on the shape of the structuring function and of its size. The details that are

smaller than the structuring function are lost and the reconstructed function is a smoothed version of the original. The subtraction of components for updating the residue may generate fragments of the signal that are smaller than the structuring function. In the continuous case, these fragments are described by decreasing the size of the structuring function (homothetics with $k < 1$). For the discrete case, these fragments increase the error of the representation. For noisy signals, the performance of the morphological signal representation depends on the probability density function of the noise as well. Since the support of the components are skeleton subsets, morphological representation is expected to inherit the poor performance of skeleton in presence of noise (lower semicontinuous transform [5]).

The main difference between binary and graylevel representation is in the definition of the components as simple functions. In the binary case, their support functions are curves with constant value. If the structuring function is a disk, the components are Blum ribbons and the functions l_i represent their spines. For graylevel signals and images, the components are simple in the sense that the opening is a simplified (filtered) version of the original. Their support functions are multivalued functions defined over a compact support. Representations in series of single-valued components can be derived by using slightly modified constructive transforms. One possibility is to select only the centers x which have the maximum value on the loci function, i.e., $\gamma(r) = [\max_x(r \ominus \kappa g^s)(x)] \oplus \kappa g$. Another possibility is to threshold the loci functions: $\gamma(r) = T_h(r \ominus \kappa g^s) \oplus \kappa g$, where $T_h(x) = 1$ if $x \geq h$ and $T_h(x) = 0$ otherwise. Both constructive transforms give qualitatively similar results as the ones obtained by multiscale opening. For example, the final representations of the test signal (Fig. 1) by using single-valued components have PSNR's of 37 dB (maximum value of loci) and of 34.65 dB (thresholded loci for $h = 1$), respectively.

The components derived so far are overlapped scaled structuring functions. Their support is compact. The number of centers of each component may range from several centers to hundreds and even thousands. Most of these centers are redundant or have a small effect in the reconstruction. Since the l_i functions are skeleton subsets of the current residue, minimal skeleton approaches can be used for redundancy reduction [6]. A very simple approach, proven to give good results, is to select the centers of the nonoverlapping structuring function fitted beneath the current residue for each component in a certain order, e.g., in lexicographic order. The constructive transform is: $\gamma(r) = [\min \parallel \text{supp}(r \ominus \kappa g^s) \parallel] \oplus \kappa g$, where supp

is the support function and min$\|\ .\ \|$ is computed as the minimum according to the considered order. The components become homothetics of the maximal structuring function. The procedure eliminates the redundancy due to the overlap of the structuring functions but does not guarantee a minimal number of components. The method is also sensitive to the ordering principle used. The quality of the reconstruction for discretized and single-valued support is good (PSNR=33.5 dB for the test signal) with the advantage of a reduction in the size of the representation.

Significant reduction in the size of the representations is achieved by using the discretized single-valued representations derived for a library of structuring functions, let it be $S = \{g_1, \ldots, g_N\}$. Thus, parts of the signal/image can be better described. Each structuring function g_j generates its component $\gamma_j(r_i)$ for the current residue r_i. The constructive transform γ_j that generates the best component f_i (e.g.,in L_2 sense) is considered: $\gamma(r_i) = \{\gamma_j(r_i) | \min_j L_2(R_{\{r_i,(r_i)_{\kappa_i g_j}\}})\}$. We have heuristically developed a library of 32 structuring functions, inspired by the features that exist in natural images (plateaux, slanted regions, edges), to be used for image decomposition. Some of these structuring functions are shown in Fig. 2. Partial reconstructions of the test image Lenna of size 512×512 by using the multiple structuring functions transform with discrete single-valued components are presented in Fig. 3. High compression ratio's ($50:1 - 100:1$) are obtained at a quite good quality of the reconstructed image. The compression ratio's have been evaluated by using Huffman coding for the labels of the structuring functions and their corresponding scales and RLC followed by Huffman coding for the coordinates of the components.

4 Adaptive schemes

Since the quality and the compactness of the representation is determined by the match between the structuring function and the signal, the definition of adaptive decompositions is natural. The task of a real adaptive decomposition is very complex. In the sequel we briefly discuss two simplified schemes, one adapting the structuring function to the signal and another one adapting the size of the signal in order to speed up the decomposition procedure.

The first scheme aims at finding some structuring functions such that they fit to the family of residues after multiple dilations. The problem can be simplified to a binary setting instead of a graylevel one

by observing that, for any threshold, once the dilated structuring function is within the residue, then the thresholded dilated structuring function is included in the thresholded residue. Furthermore, when flat structuring functions are considered, their shape can be fully determined from a single thresholded residue. Thus, the idea is to consider a reduced set of flat structuring functions and to adapt their shape. For parallelepiped structuring functions this means the computation of the height h and width l of a structuring function g that, being dilated n times, still fits in the residue. The width L and the height H of the dilated structuring functions are:

$$H = v + nh,\ L = v + n(l-1) \tag{3}$$

where v is the value of the locus function, e.g., $v = 1$. The adaptive decomposition iteratively determines an optimal threshold H. The threshold could be the global one that gives the best approximation of the signal (e.g., in L_2 norm). Furthermore, the maximal component is extracted from the binarized residue $T_H(r)$ similar to the extraction of the first component in MSD for binary images [2]. The size of the maximal component gives the maximal width L. The height and the width of the structuring function g are computed by finding integer solutions for equations (3). Generally, an exact solution does not exist and the best solution that approximates L and H must be selected. (3) can be solved by exhaustive search. The number of solutions may be minimized by considering only a small number of admissible values for l and h. Finally, the residue is updated. If the residue is empty, or the reconstruction error is less then a predefined value, the procedure ends, otherwise it reiterates for a new selected threshold H. The method is easy to adjust to other flat structuring functions by modifying (3).

The residual decomposition procedure ends when the residue is empty or its measure is less than a predefined threshold. It is straightforward to assume that the measure of the initial signal $L(f)$ (e.g., L_2 norm or the volume beneath the signal "surface") is a major parameter that determines the number of components, i.e., the compactness of the representation. It is very likely that, for two signals, f and g, with $L(f) < L(g)$, the number of components for the representation of f is smaller than the one for the representation of g. Thus, the second scheme is based on the idea of modifying f before the decomposition, so that the measure of the transformed signal in the L sense is less than the measure

of the original one. The representation is performed for the modified signal and an inverse transform is applied to recover the original one. It is obvious that not only transform must yield a significant reduction in the measure of the signal, but also the direct and the inverse transforms must be computationally fast. Let us consider the following simple transform that splits the signal f in two components, f_+ and f_- :

$$f_+ = \begin{cases} f(x) & \text{if } f(x) \geq T \\ 0 & \text{otherwise} \end{cases}, \; f_- = \begin{cases} T - f(x) & \text{if } f(x) < T \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where T is a suitable value within the range of f. f is reconstructed as $f = T + f_+ - f_-$. Relation (4) can be seen as the approximation of the original signal f by a constant signal T. The approximation error is split in (f_+, f_-) in order to further operate on positively defined functions. We are interested in the value of T that minimizes the measure of f_+ and f_-. Convenient selections for T are: the arithmetic mean of f, the median of f, etc. The computation of (4) is very fast and the reconstruction is very simple. The decomposition of f in f_+ and f_- may be regarded as the adaptation of the size of the image to the decomposition scheme. Since the transform is very simple, we consider an adaptive algorithm that, by using (4), continuously updates the current residue. We mention that only the last T_j value is needed for the j-component reconstruction:

$$f_j = T_j + \sum_{i=1}^{j} f_{i+} - \sum_{i=1}^{j} f_{i-} \tag{5}$$

The representation is divided in two sets of components: one set represents f_+, the other represents f_-.

The adaptive methods performs well. PSNRs of 34.3 dB and 33.15 dB respectively, have been achieved for the test signal of Fig. 1. Very good partial reconstructions are obtained for a small number of components, i.e., a PSNR of 25 dB is obtained for only 32 component reconstructions for both methods and a PSNR of 30 dB for 72 components (adaptive structuring functions), 84 components (adaptive threshold), respectively.

5 Conclusions

A general approach for signal and image representations is developed. By using a constructive transform, simple components are extracted and signals are represented as a sum of their simple components. Constructive transforms based on mathematical morphology are analyzed and adaptive decomposition schemes are investigated. The derived representations encode efficiently the signal and provide signal descriptions at varying levels of detail. The experimental results have shown that residual representations are efficient in lossy image compression.

References

[1] P. Maragos, "Pattern Spectrum and Multiscale Representation", *IEEE Trans. on PAMI*, vol. 11, no. 7, pp. 701-716, July 1989.

[2] I. Pitas, A.N. Venetsanopoulos, "Morphological Shape Decomposition", *IEEE Trans. on PAMI*, Vol. 12, No. 1. pp.38-45, 1990.

[3] I. Pitas, "Morphological Signal Decomposition", *Proc. of the IEEE Intl. Conf on ASSP*, pp. 2169-2172, Albuquerque, 1990.

[4] I. Pitas, A. Maglara, "Range Image Analysis by Using Morphological Signal Decomposition", *Pattern Recognition*, Vol. 24, No. 2, pp. 165-181, 1991.

[5] J. Serra, *Image Analysis and Mathematical Morphology*, Academic Press, New York, 1982.

[6] P. Maragos, R. W. Schaffer, "Morphological skeleton representation and coding of binary images", *IEEE Trans. on ASSP*, ASSP-34(5), pp. 1228-1244, Oct. 1986.

[7] D. Coltuc, I. Pitas, "Morphological Signal Decomposition", *IEEE Workshop on Nonlinear Signal and Image Processing*, Neos Marmaras, Greece, pp.983-986, 1995.