NONLINEAR SIGNAL PROCESSING FILTERS:

A UNIFICATION APPROACH

L. Yaroslavsky,
Dept. of Interdisciplinary Studies, Faculty of Engineering,
Tel Aviv University, Tel Aviv 69978, Israel,
Phone: 972-3-640-7366, Fax: 972-3-641-0189
e-mail: yaro@eng.tau.ac.il

ABSTRACT

The paper addresses the problem of classification of nonlinear filters for digital signal processing and of unification of approaches to their analysis, design and implementation. It is assumed that the filters work in a moving window and, in each position of the window, produce an output signal sample as an estimate obtained from input signal samples within the window. The keys of the classification and unification are notions of estimation and neighborhood building operations. On the base of thorough analysis of a large variety of known filters, a set of estimation and multi stage (in general) neighborhood building operations is introduced that can serve as building blocks for the filter design. In conclusion, application of the approach to the filter design and implementations is briefly discussed.

1. INTRODUCTION

An approach to unification of nonlinear filters is presented that allows to:

- navigate in the ocean of nonlinear filtering methods;
- reveal interrelations between different methods;
- provide a common language for researches in nonlinear filers;
- provide a common base to synthesis and optimization of nonlinear filters.
- facilitate association with computer and hardware implementation;

The following main assumptions constitute the base of the approach:

- Filters work in a moving window.
- In each position $\boldsymbol{k}$ of the window, filters generate from the window samples $\{b_n\}, n = 1,2,..,Window\ size$ an output signal sample $\hat{a}_k$ that is regarded as an estimate of input signal "true" value a_k in this position.
- The estimate is obtained by means of a certain estimation operation ***ESTM*** applied to a certain set $NBH(a_k)$ of values that constitute the sample's neighborhood:

$$\{b_n\} \rightarrow \hat{a}_k : \hat{a}_k = ESTM(NBH(a_k))$$

- The neighborhood is derived in a certain way from the window samples.
- Filters can be specified in terms of the method of building neighborhood and of the type of the estimation operation.

Therefore, key notions in the classification are those of neighborhood, neighborhood building and estimation operations. A set of estimation operations and neighborhood building operations is suggested in the paper on the base of a thorough analysis of different nonlinear filters published in the literature ([1-4]). Iterative, cascade and recursive implementations of the filters as well as other possible generalizations and immediate implications of the suggested approach to filter analysis, design and efficient soft and hardware implementation are also discussed.

2. NOTION OF THE NEIGHBORHOOD. W-NEIGHBORHOOD

Neighborhood is a set of elements and attributes that are inputs to the estimation operation. The process of forming neighborhood can be, in general, multi stage, or hierarchical. We shall denote *n*-stage neighborhoods by NBH^n. An important special case is that of the one-stage neighborhood. On the one-stage process, all window elements form the neighborhood and one has only to compute their attributes necessary for a particular estimation operation. We shall call this primary window neighborhood ***Wnbh***: $NBH^1 \equiv Wnbh$.

Neighborhood attributes may be associated both with elements of the neighborhood and with the neighborhood as a whole. The following attributes may be used as attributes of neighborhood elements:

- **Value**.
- **Co-ordinate** in the window.
- **Rank**, or position of the element in the variational row formed from all elements of the neighborhood by sorting them from minimum to maximum. Rank of an element shows the number of elements of the neighborhood that have lower value than that of the given element.

- **Cardinality**, or number of neighborhood elements with the same value (applicable for quantized vales).
- **Geometrical features**, such as gradient, curvature, etc.

Attributes **Rank** and **Cardinality** are interrelated and can actually be regarded as two implementations of the same quality. While **Rank** is associated with variational row, **Cardinality** is associated with histogram over the neighborhood. Choice between them is governed, in particular, by their computational complexity. The computational complexity of operating with histogram is ***O(Q)***, where ***Q*** is the number of quantization levels used for representing signal values. Computational complexity of operating with variational row is ***O(N)***, where ***N*** is the number of neighborhood elements. Histogram of the window elements in each window position can be easily computed recursively from the histogram in the window previous position with only ***O(1)*** operations. Recursive formation of the variational row is also possible with the complexity of ***O(N)*** operations.

Attributes of neighborhood as a whole are usually used in multi stage process of building neighborhoods. As such attributes, results of any estimation operation can be used.

3. ESTIMATION OPERATIONS

Estimation operations used to generate filter current sample output value are a many-to-one mappings of neighborhood attributes to one output value. Selection of a particular estimation operation is governed by a priori knowledge regarding signal properties and the criterion of the processing quality. Here is a list of typical estimation operations composed on the base of analysis of nonlinear filters known from the literature.

- **MEAN(*NBH*)**: arithmetic mean of elements of the neighborhood. As it is well known, **MEAN** is an optimal MAP- (Maximal A Posteriori Probability) estimation of a location parameter of data in the assumption that data are observations of a single value distorted by an additive noncorrelated Gaussian random values (noise).
- **PROD(*NBH*)**: product of elements of the neighborhood. Product is an operation homomorphic to the addition involved in **MEAN** operation: sum of logarithms of a set of values is logarithm of their product.
- **K_ROS(*NBH*)**: ***K***-th rank order statistics over the neighborhood. The operation picks up the element that occupies ***K***-th place (has *rank* ***K***) in the variational row. Special cases of **K_ROS(*NBH*)** operation are:
- **MIN(*NBH*)**: minimum over the neighborhood;
- **MAX(*NBH*)**: maximum over the neighborhood;
- **MEDN(*NBH*)**: median over the neighborhood, that is central element of the variational row.

  These operations are optimal MAP estimations for other then additive Gaussian noise models. For instance, if neighborhood elements are observations of a constant distorted by addition to it independent random values with exponential distribution density, **MEDN(*NBH*)** is known to be the optimal MAP estimation of the constant. If additive noise samples have one-sided distribution and affect not all data, **MIN(*NBH*)** or **MAX(*NBH*)** might be optimal estimations. One can imagine models for which intermediate Rank Order Statistics may be optimal estimations.
- **MODE(*NBH*).** This estimation operation provides the value of the neighborhood element with the highest cardinality. It is sort of an analogue to **MAX(*NBH*).**
- **RAND(*NBH*)**. This estimation produces a pseudo-random number taken from an ensemble with the same probability distribution as that of elements of the neighborhood. This operation may look exotic and is not a traditional one in the theory of estimation. However it has a remarkable property that it generates an estimation that is statistically (in terms of the distribution density) akin to the neighborhood elements: it has the same mean value, standard deviation, all other moments, etc. In certain cases it may serve as a sort of a "soft" alternative to the above "hard" estimations ([4]).
- **SIZE(*NBH*)**. This operation provides the number of elements of the neighborhood.
- **STDEV(*NBH*)**: standard deviation over the neighborhood. Conventional standard deviation operation for evaluation of data variance is nonrobust agains outliers. In case of outliers present in the data, Interquantil Distance **IQDIST(*NBH*)** may be preferable:

$$\mathbf{IQDIST}(\boldsymbol{NBH}) = \mathbf{R\_ROS}(\boldsymbol{NBH}) - \mathbf{L\_ROS}(\boldsymbol{NBH}),$$

where $\mathbf{1} \le \boldsymbol{L} < \boldsymbol{R} \le \mathbf{SIZE}(\boldsymbol{NBH})$.

One should also introduce a number of auxiliary relation operations between individual elements of the neighborhood and the entire neighborhood such as:

- **MEMB(*NBH,a*)**. This is a binary operation that evaluates by 0 and 1 membership of a certain window element **a** in the neighborhood ***NBH***.
- **DEV(*NBH,a*)** measures differences between elements of the neighborhood and certain value ***a*** obtained usually also by some estimation operation over the neighborhood. In particular, it can be used for obtaining most robust estimation **MAD** (median of absolute deviation) of dispersion of neighborhood samples ([1], 194) :

$$\mathbf{MAD}(\boldsymbol{NBH}) = \mathbf{1.483MEDN}(\mathbf{DEV}(\boldsymbol{NBH}, \mathbf{MEDN}(\boldsymbol{NBH})))$$

- **ATTR(*NBH,a*)**. This operation provides certain attribute of element **a** in the neighborhood defined by its another attribute. Examples of this operation are:
- **RANK(*NBH*,$\boldsymbol{a_k}$)**: rank or value of the element defined by its co-ordinate ***k*** in the window. As it was mentioned, ranks of neighborhood elements can be found from histogram over the neighborhood. In ([3,4]), a natural generalization was described to

ordering through the histogram: prior to ordering, histogram values are modified by some function:

$$f\_histogram = f(histogram).$$

Modified in this sense **RANK(*NBH*,a_k)** we shall denote ***F*_RANK(*NBH*,a_k).** One of the useful special cases is that of ***P***-th law nonlinearity:

$$P\_histogram = (histogram)^P .$$

In this case, ***F*_RANK(*NBH*,a_k)** will be denoted ***P*_RANK(*NBH*,a_k).**

- **COORD(NBH,a_r):** co-ordinate of the element with a certain rank **r**.

In addition note that different combinations of the basic operations can also be used for generating filter output.

4. MULTI STAGE NEIGHBORHOODS

In multi stage neighborhood forming, elements of the neighborhood formed on a previous stage are first grouped into a number of subgroups. Each of the subgroups is then treated individually for forming from them sub-neighborhoods to which intermediate estimation operations are then applied to produce a new set of values until, on the last stage, only one value is obtained as a filter output.

Selection of elements for sub neighborhoods can be based on their attributes. We shall distinguish here sub-neighborhoods ***SubWnbh*** formed according to element coordinates and sub-neighborhoods ***SubRnbh*** formed according to element ranks, although the use of other attributes for forming sub-neighborhoods is, in principle, possible as well. In the multi stage neighborhood building process, ***W***-neighborhood is formed first: for signal samples within a window centered at a certain current signal sample their attributes are found. Then elements of the W-neighborhood are grouped to form SubW-neighborhoods to which intermediate estimation operations are applied. At the result, a new neighborhood is formed which, for the 2-stage process, is final one used as input to final estimation operation that generates filter output for the current window position.

5. NEIGHBORHOOD BUILDING OPERATIONS

As neighborhood building operations, all above described "many-to-one" operations can be used. There is, however, a number of additional classes of "many-to-many" operations proved being useful for building intermediate neighborhoods in multi stage process. These are:

- **GROUP_A(*NBH*)** – attribute based formation of groups of "sub neighborhoods"

GROUP_A(*NBH*) ={***Subnbh***$_1$,***Subnbh***$_1$, …,***Subnbh***$_n$};

$$\bigcup_k Subnbh_k = NBH$$

- **SELECT_A({*Subnbh*$_k$})** – sub neighborhood attribute controlled selection of one subneighborhood from a set. There should be mentioned several special cases of sub neighborhoods that result from such a selection:

- ***SHnbh***-, or Shape-neighborhoods that are obtained from ***W***-neighborhood by applying to its elements binary weight coefficients according to their co-ordinates. In 2-D and multi-dimensional case, this corresponds to forming spatial neighborhoods of a certain shape.
- ***Qnbh***-, or Quantil-neighborhoods that are obtained from ***W***-neighborhood by selection from it elements (order statistics) whose ranks $\{R_r\}$ satisfy inequality
  $$1 < R_{left} < R_r < R_{right} < SIZE(Wnbh)$$
  where R_{left}, R_{right} are boundaries of an interval in the variational row.
- $EVnbh(NBH; a_k; \varepsilon Vpl; \varepsilon Vmn)$, or "epsilon-V"-neighborhood of element a_k. $EVnbh(a_k)$ is a subset of elements with values $\{a_n\}$ that satisfy inequality:
  $$a_k - \varepsilon Vmn \le a_n \le a_k + \varepsilon Vpl .$$
- $KNVnbh(NBH; a_k, K)$, or "***K*** nearest by value"-neighborhood of element a_k. $KNVnbh(NBH; a_k, K)$ is a subset of ***K*** elements with values $\{a_n\}$ closest to that of element a_k.
- $ERnbh(NBH; a_k; \varepsilon Rpl : \varepsilon Rmn)$

  and $KNRnbh(NBH; a_k, K)$-neighborhoods that are analogs of above two "**V**-neighborhoods" but formulated in terms of element ranks rather then of their values.
- $CLnbh(NBH; a_k)$, or "Cluster" neighborhood of element a_k. It is formed from all neighborhood elements that belong to the same cluster of neighborhood's histogram as that of element a_k.
- $FLATnbh(NBH; Thr)$. It is formed from all neighborhood elements with gradient (or similar geometrical measure of signal/image "flatness") is lower than a certain threshold Thr.
- **FUNC(*NBH*)** is element wise functional transformation of neighborhood elements. Its special cases are
- **MULT_ATTR(*NBH*)** (multiplying elements of the neighborhood by some weights) and **REPL_ATTR(*NBH*)** (replicating them certain number of times). The weight coefficients and replication factors are defined by certain attributes of the elements. One can distinguish following versions of these neighborhood building operations (for brevity, only version of **MULT_ATTR** will be listed):
  - **MULT_C(*NBH*):** multiplication by weighting coefficients defined by element co-ordinates,

- **MULT_V(*NBH*):** multiplication by weighting coefficients defined by element values,
- **MULT_R(*NBH*):** multiplication by weighting coefficients defined by element ranks,
- **MULT_H(*NBH*):** multiplication by weighting coefficients defined by cardinality of the neighborhood elements,
- **MULT_G(*NBH*):** multiplication by weighting coefficients that are defined by geometrical attributes of the neighborhood elements.
- **MULT_CombA(*NBH*)**: weight coefficients depend on combination of attributes , for instance, from both co-ordinates and ranks of neighborhood elements (**MULT_CR(*NBH*)).**

On intermediate stages of multi stage neighborhood formation, estimation operations can be applied not only element wise but to intermediate sub neighborhoods regarded in terms of their global attributes. In order to avoid unnecessary detalization, we'll give only two examples of such a sub neighborhood wise operation:

$\mathbf{NBH^2 = MIN\_Std}(\boldsymbol{SubWnbh_1, SubWnbh_2, ..., SubWnbh_n})$ that selects from a set of sub neighborhoods one that has minimal standard deviation and

$\mathbf{NBH^2 = MIN\_RNG}(\boldsymbol{SubWnbh_1, SubWnbh_2, ,,,, SubWnbh_n}$ that selects a set of sub neighborhoods one that has minimal range **RNG=MAX(*NBH*)-MIN(*NBH*).** Obviously, this list can be extended further.

6. ITERATIVE, CASCADE AND RECURSIVE FILTERING

The theory based on the idea of signal local approximation ([3]) says that neighborhood needed for finding an estimation of the window central pixel depends, in general, on its true value which is unknown and has to be estimated. This implies that optimal estimation algorithm should be, in principle, iterative:

$$\hat{a}_k^{(t)} = \mathbf{ESTM}\left(NBH\left(a_k^{(t-1)}\right)\right).$$

Iterative computation is also a way to implement estimation operations that are defined not explicitly but rather by a certain optimization criterion, such as, for instance, in M-filters ([1], p. 122).

An important problem of the iterative filtering is that of adjustment of neighborhood building and estimation operations according to changing statistics of noise that takes place in course of iterations. This may require iterative wise change of the filter parameters or even its entire structure. One of the possible solutions of the problem is to combine in one filter two filters and switch between them under control of a certain third filter that evaluates the changing statistics.

A variety of iterative filtering is cascade filtering when the estimation operation is also modified in course of iterations.

Computational expenses associated with iterative and cascade filtering can be reduced by using , in the process of scanning signal by the filtering window, as window samples those that are already estimated in previous positions of the window. This implementation is known as recursive

7. CONCLUSION AND DISCUSSION

One can show that all nonlinear filters known from the literature ([1-4]) are built on a base of a finite set of estimation and neighborhood building operations introduced in Sects. 3-5. In this sense, treatment of nonlinear filters in terms of estimation and neighborhood building operation does provide a unified framework for the representation of signal filters and reveals relationship between different methods. Within this framework, it is also almost straightforward to suggest many new modifications of filters. The set of estimation and neighborhood building operations introduced may serve as a base for the development of program packages and programmable VLSI chips for efficient software and hardware implementation of nonlinear filters for signal and image processing. In addition, analysis of neighborhood formation in Sect. 4 (Fig. 1) clearly reveals an important common feature of the filters: they are inherently parallel and can therefore naturally be implemented in a form of multi layer neuro-morphic networks.

The author believes that the presented approach offers also a way for the filter synthesis and optimization. The approach assumes that the filter design and optimization for a given quality measure is subdivided into several neighborhood building and estimation steps. Correspondingly, the quality criterion should also be subdivided to allow optimization of these steps. This can be an alternative to a traditional signal model based approach to the filter design when the appropriate model involves too many parameters to allow direct filter mapping optimization.

REFERENCES:

1. J. Astola, P. Kuosmanen, Fundamentals of Nonlinear Digital Filtrering, CRC Press, Roca Baton, New York, 1997
2. I. Pitas, A. N. Venetsanopoulos, Nonlinear Digital Filters. Principles and Applications. Kluwer, 1990
3. L. Yaroslavsky, M. Eden, Fundamentals of Digital Optics, Birkhauser, Boston, 1996
4. V. Kim, L. Yaroslavsky, Rank algorithms for picture processing, Computer Vision, Graphics and Image Processing, v. 35, 1986, p. 234-258