

MAX-MIN DIFFERENCE EQUATIONS AND RECURSIVE MORPHOLOGICAL SYSTEMS

Petros Maragos

Division of Applied Sciences, Harvard University, Cambridge, MA 02138, USA

ABSTRACT

In this paper we introduce a theory for recursive systems modeled via max/min difference equations. Their applications include fast algorithms for distance transforms and signal envelope estimation. Such systems correspond to dilations or erosions by infinite-support structuring functions. For their understanding we introduce the class of dilation (or erosion) translation invariant systems along with several concepts and tools for their analysis: causality, stability, eigen-functions, a transform that quantifies the slope content of signals, and a transfer function.

1. INTRODUCTION

All morphological systems processing signals are based on parallel or serial interconnections of morphological dilations \oplus or erosions \ominus [6, 7, 3]

$$\begin{aligned} x(n) \oplus g(n) &= \bigvee_k x(k) + g(n - k) \\ x(n) \ominus g(n) &= \bigwedge_k x(k) - g(k - n) \end{aligned}$$

where \bigvee =supremum (or maximum if the set of indexes k is finite) and \bigwedge = infimum (or minimum). The vast majority of theory and applications of morphological signal processing assumes that, for discrete-variable input signals $x(n)$, the structuring function $g(n)$ has a finite support and hence the above moving max/min of additions takes place only over a finite window of input samples. There is, however, an important application that requires max/min operations by recursing on output samples. This occurs during the computation of the distance transform of binary images [5, 1], which is useful in its own right e.g. for image segmentation via watersheds, but it can also provide much additional information such as the skeleton, multiscale

erosions, and granulometries [2]. In addition, in the context of recursive rank-order filtering which has been studied to a limited extent [4], recursive max and min operations are special cases of recursive rank filters but also the most important since any rank-order operation is a minimum of maxima or maximum of minima [3].

In this paper we introduce a theory for recursive dilations and erosions (acting on discrete-variable signals) by modeling them via nonlinear difference equations of the max type

$$y(n) = \left(\bigvee_{k=1}^K a_k + y(n - k) \right) \vee \bigvee_{m=0}^M b_m + x(n - m) \quad (1)$$

or the min type. Whenever such an equation has a recursive part, we show that this corresponds to dilating the input signal with an infinite-support structuring function. Further, we derive many interesting results concerning the impulse response, causality, stability, and eigen-functions of such systems by drawing analogies with similar properties of linear systems described by linear difference equations. To aid the analysis we introduce a transform of signals that quantifies their 'slope content' and derive a transfer function for dilation or erosion systems related to their eigen-functions. Finally, we present numerical examples and applications to estimation of signal envelopes.

2. MAX DIFFERENCE EQUATIONS

Eq. (1) describes a nonlinear system Ψ with input the discrete-time signal x and output $y = \Psi(x)$. All signals in this section are defined on \mathbb{Z} and their range is $\mathbb{R} \cup \{-\infty\}$, where \mathbb{R} and \mathbb{Z} are the sets of reals and integers. The **support** of any signal $x(n)$ is defined by $\text{Spt}(x) \triangleq \{n \in \mathbb{Z} : x(n) > -\infty\}$. In (1) assume that all a_k and b_m are in $\mathbb{R} \cup \{-\infty\}$; if $a_k = -\infty$ the term with $y(n - k)$ is not used in the difference equation. K is the order of the equation, assuming $a_K > -\infty$. To

solve (1) in forward time $n \geq n_0$ we need K **initial conditions** $IC(n_0)$, where

$$IC(n) \triangleq \{y(n-1), \dots, y(n-K)\}$$

If all the values in $IC(n_0)$ are $-\infty$, the initial state of the system does not affect its response.

To analyze (1) and other similar morphological systems Ψ we introduce the following concepts. First, two useful basic signals are the **zero impulse** δ and **zero step** s :

$$\delta(n) \triangleq \begin{cases} 0, & n = 0 \\ -\infty, & n \neq 0 \end{cases} ; \quad s(n) \triangleq \begin{cases} 0, & n \geq 0 \\ -\infty, & n < 0 \end{cases}$$

Now we define the **impulse response** $g = \Psi(\delta)$ as the system's response when the input is the impulse and $IC(0) = -\infty$. A system Ψ is **causal** if its defining rule at each time instant depends only on present and/or past input values and possibly on past output values, but not on future input or output values. A system is **stable** if a bounded (within its support) input signal yields a bounded output signal (within its support).

Consider the **1st-order** max difference equation, with $a, b \in \mathbb{R}$,

$$y(n) = [y(n-1) + a] \vee [x(n) + b] \quad (2)$$

Assuming causality and $x(n) = -\infty$ for all $n < 0$, by induction on $n \geq 0$ we find its solution

$$y(n) = \underbrace{\left(b + \bigvee_{k=0}^n x(k) + (n-k)a \right)}_{= x \oplus g(n)} \vee [an + a + y(-1)]$$

where $g(n)$ is the impulse response

$$g(n) = an + b + s(n) = \begin{cases} an + b, & n \geq 0 \\ -\infty, & n < 0 \end{cases}$$

Namely, $g(n)$ is the solution $y(n)$ of (2) when $x(n) = \delta(n)$ and $y(-1) = -\infty$. Thus the general solution $y(n)$ of (2) is the maximum of the $(-\infty)$ -state response (i.e., the dilation $x \oplus g$) and the $(-\infty)$ -input response due only to the initial condition $y(-1)$. The system is stable only if $a = 0$. Similar results are also true for the general K^{th} -order max difference equation. Initial conditions $\neq -\infty$ could be useful in some applications; e.g., if $y(-1) > -\infty$, the solution of (2) is constrained to be $\geq an + a + y(-1)$. However, in the rest of the paper we shall assume $-\infty$ initial conditions.

DTI Systems

A system Ψ is called **dilation translation invariant (DTI)** iff it distributes over any supremum of input signals, i.e., $\Psi(\bigvee_i x_i) = \bigvee_i \Psi(x_i)$, and is translation-invariant, i.e., $\Psi[x(n-k) + c] = c + [\Psi(x)](n-k)$. Since each input signal can be represented as

$$x(n) = \bigvee_{k=-\infty}^{\infty} x(k) + \delta(n-k) = \bigwedge_{k=-\infty}^{\infty} x(k) - \delta(n-k)$$

a system Ψ is DTI iff it is a morphological dilation by its impulse response g . The following shows that the impulse response uniquely characterizes a DTI system and determines many of its properties.

Theorem A:¹ A system Ψ with $g = \Psi(\delta)$ is

- (i) DTI iff $\Psi(x) = x \oplus g$.
- (ii) Causal iff $g(n) = -\infty \forall n < 0$.
- (iii) Anti-Causal iff $g(n) = -\infty \forall n > 0$.
- (iv) Stable iff $\sup\{|g(n)| : n \in \text{Spt}(g)\} < \infty$. \square

The affine signals $x(n) = an + b$ are **eigen functions** of DTI systems because the corresponding outputs are $y(n) = an + b + G(a)$ where $G(a) = \bigvee_n g(n) - an$ is the corresponding eigen-value. Viewing $G(a)$ as a transform for the signal $g(n)$ with variable the **slope** a , we define the following transform² $\mathcal{A} : x(n) \mapsto X(a)$ for any signal $x(n)$:

$$X(a) \triangleq \bigvee_n x(n) - an, \quad a \in \mathbb{R}$$

Some properties of the transform \mathcal{A} include

$$y(n) = x(n) \oplus g(n) \implies Y(a) = X(a) + G(a)$$

and others summarized in Table 1. Thus $G(a)$ plays the role of a ‘transfer function’ (actually a **slope response**) of the DTI system. We have also found an ‘inverse’ transform $\mathcal{A}^{-1} : X(a) \mapsto \hat{x}(n)$ which yields a signal

$$\hat{x}(n) \triangleq \bigwedge_{a \in \mathbb{R}} X(a) + an \geq x(n)$$

that is sometimes equal to $x(n)$ and never smaller. Specifically, at any time instant n , the reconstructed

¹Due to the workshop's limitation in the number of pages, we do not include the proofs of our results; they will be available in a forthcoming longer version.

²The definition of \mathcal{A} for continuous-time signals $x(t)$ is identical: $X(a) = \bigvee_t x(t) - at$; likewise for \mathcal{A}^{-1} . All the properties in Table 1 also apply to the continuous-time transform.

signal $\hat{x}(n)$ is equal to the original signal $x(n)$ iff

$$x(n) \geq \frac{px(n-q) + qx(n+p)}{p+q} \quad \forall p, q > 0 \quad (3)$$

[Given a function $f : D \rightarrow \mathbb{R}$, where $D \subseteq \mathbb{R}$ or $D \subseteq \mathbb{Z}$, f is **concave** iff f satisfies (3) for all $n \in D$; f is **convex** if the \geq in (3) is replaced by \leq $\forall n$.] Now note that the transform $X(a)$ is always a convex function, and the reconstructed signal $\hat{x}(n)$ is always concave. Thus, the ‘inverse’ transform \mathcal{A}^{-1} yields a concave signal $\hat{x}(n) \geq x(n) \forall n$. Actually, the signal $\hat{x}(n)$ is the smallest concave **upper envelope** of $x(n)$; hence it can be constructed as the piecewise linear interpolation of the points where $x(n) = \hat{x}(n)$, i.e., points where (3) is true.

Examples (see also Table 1 and Fig. 1): If $x(n) = 0$ for $|n| \leq N$ and $-\infty$ else, then $X(a) = N|a|$ and $x = \hat{x}$. Now if $y(n) = 0$ for $n = \pm N$ and $-\infty$ else, then $Y(a) = X(a)$ and $\hat{y}(n) = y(n)$ for $|n| \geq N$, but $\hat{y}(n) > y(n)$ for $|n| < N$.

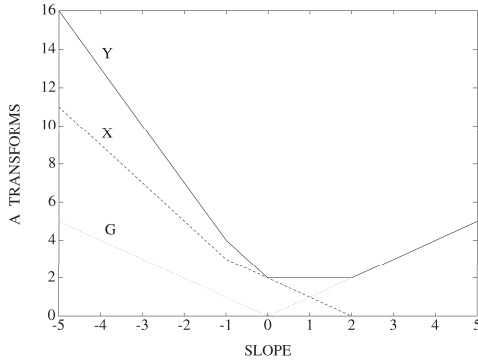


Figure 1: \mathcal{A} transforms $X(a)$, $G(a)$, and $Y(a) = X(a) + G(a)$ of the signals $x(n) = 0, 2, 1$ for $n = 0, 1, 2$, $g(n) = 0$ for $n = -1, 0, 1$, and $y(n) = x \oplus g(n)$. Note: $\mathcal{A}^{-1}[Y]$ yields $\hat{y}(n) = y(n) = 0, 2, 2, 2, 1$ for $n = -1, 0, 1, 2, 3$ and $-\infty$ else.

Concluding, the transform \mathcal{A} quantifies the ‘slope content’ in signals in a similar way as the Fourier transform quantifies the ‘frequency content’ of signals. If $g(n)$ is the impulse response of a DTI system, then its transform $G(a)$ acts as a ‘slope response’ because it is the (additive) eigen-value for affine input eigen-signals $x(n) = an + b$. Thus $G(a)$ is conceptually similar to the frequency response of a linear time-invariant system which is the (multiplicative) eigen-value for exponential input eigen-signals. In addition, the dilation of two concave signals can be done by adding their \mathcal{A} transforms (see Fig. 1).

TABLE 1: Properties & Examples of \mathcal{A}

Signal $x(n)$	Transform $X(a)$
$[b + x(n)] \vee [c + y(n)]$	$[b + X(a)] \vee [c + Y(a)]$
$x(n - k)$	$X(a) - ka$
$x(n) + a_0 n$	$X(a - a_0)$
$x(n) \oplus y(n)$	$X(a) + Y(a)$
$rx(n)$, $r > 0$	$rX(a/r)$
$x(-n)$	$X(-a)$
$x(n) \leq y(n) \quad \forall n$	$X(a) \leq Y(a) \quad \forall a$
$\bigvee_n x(n) = X(0)$	$\bigwedge_a X(a) \geq x(0)$
$x(n) \wedge y(n)$	$\leq X(a) \wedge Y(a)$
$x(n) + y(n)$, concave x or y	$\leq \bigwedge_b X(b) + Y(a - b)$
$\delta(n)$	0
$a_0 n$	$-\delta(a - a_0)$
$a_0 n + s(n)$	$-s(a - a_0)$
$a_0 n + s(-n)$	$-s(a_0 - a)$

Recursive DTI Systems

Theorem B: The max difference equation (1) corresponds to a causal DTI system iff (i) whenever $x(n) = -\infty$ for all $n < n_0$ then $y(n) = -\infty$ for all $n < n_0$, where n_0 is an arbitrary but otherwise fixed time instant, and (ii) the required initial conditions $IC(n_0)$ are $-\infty$. \square

Henceforth we make the above two assumptions for systems described by (1). There are two major subclasses of such systems:

Finite Impulse Response (FIR) DTI systems, when $a_k = -\infty$ for all k . Then (1) has no recursive part, and the impulse response

$$g(n) = \begin{cases} b_n & \text{if } n = 0, 1, \dots, M \\ -\infty & \text{if } n < 0, n > M \end{cases}$$

has finite support. This class is identical with the class of all morphological dilations with finite-support structuring elements.

Infinite Impulse Response (IIR) DTI systems, when $a_k \neq -\infty$ for at least one k . The example of the 1st-order system (2) demonstrates that such systems have an impulse response g of infinite support.

We can also bring (1) to the standard form

$$y(n) = \max[y(n-1) + a_1, \dots, y(n-K) + a_K, x(n)] \quad (4)$$

by pre-dilating the input signal x with the finite-support signal b defined as $b(n) = b_n$ for $n = 0, \dots, M$ and $-\infty$ else. We also assume that $b_0 = 0$, because a nonzero b_0 only adds a constant b_0 to the output y .

1st-order system $y(n) = \max[y(n-1) + a_1, x(n)]$: Its impulse response is $g(n) = a_1 n +$

$s(n)$, and its slope response is

$$G(a) = -s(a - a_1) = \begin{cases} +\infty, & a < a_1 \\ 0, & a \geq a_1 \end{cases}$$

It acts as a ‘slope highpass’ filter since it passes from the input signal only those segments whose slopes are $\geq a_1$.

K^{th} -order system: Applying \mathcal{A} to (4) and assuming $X(a)$ is finite yields

$$G(a) = \max[G(a) - a + a_1, \dots, G(a) - Ka + a_K, 0]$$

If $m = \arg \max_k \{a_k/k\}$, the slope response is

$$G(a) = -s\left(a - \frac{a_m}{m}\right)$$

and the (reconstructed via \mathcal{A}^{-1}) envelope of the impulse response is $\hat{g}(n) = na_m/m + s(n) \geq g(n)$. Thus, although over short time periods $g(n)$ has the shape induced by the sequence $\{a_k\}$, over large time scales it behaves like its concave upper envelope $\hat{g}(n)$. Actually, if $a_k < 0$ for all $k < K$ and $a_K = 0$, the infinite impulse response $g(n)$ becomes periodic! Examples are shown in Fig. 2.

3. MIN DIFFERENCE EQUATIONS

In this section we deal with signals $x : \mathbb{Z} \rightarrow \mathbb{R} \cup \{+\infty\}$ whose support is defined as $\text{Spt}(x) \triangleq \{n : x(n) < \infty\}$. A system $\Psi : x(n) \mapsto y(n)$ is called **erosion translation-invariant (ETI)** if it distributes over any infimum of inputs, i.e., $\Psi(\bigwedge_i x_i) = \bigwedge_i \Psi(x_i)$ and is translation invariant. We focus on systems Ψ described by the following min difference equation

$$y(n) = \left(\bigwedge_{k=1}^K a_k + y(n-k) \right) \wedge \bigwedge_{m=0}^M b_m + x(n-m) \quad (5)$$

for which we henceforth assume that (i) whenever $x(n) = \infty$ for all $n < n_0$ then $y(n) = \infty$ for all $n < n_0$, where n_0 is an arbitrary but otherwise fixed time instant, and (ii) the required initial conditions $IC(n_0)$ are ∞ . This guarantees that (5) describes a causal ETI system.

If we use the signal $-\delta(n)$ as the zero impulse, then the *impulse response* of an ETI system Ψ is defined as the signal $h = \Psi(-\delta)$.

Theorem C: A system Ψ with $h = \Psi(-\delta)$ is

(i) ETI iff

$$\Psi(x)(n) = x(n) \ominus (-h(-n)) = \bigwedge_k x(k) + h(n-k).$$

(ii) Causal iff $h(n) = +\infty \forall n < 0$.

(iii) Stable iff $\sup\{|h(n)| : n \in \text{Spt}(h)\} < \infty$. \square

Let us define the signal transform $\mathcal{A}_e : x(n) \mapsto X_e(a)$ with

$$X_e(a) \triangleq \bigwedge_n x(n) - an, \quad a \in \mathbb{R}$$

Then the affine signals $x(n) = an + b$ are *eigen functions* of any ETI system Ψ because the corresponding outputs are $y(n) = an + b + H_e(a)$, where $H_e(a)$ is the corresponding eigen-value and equal to the \mathcal{A}_e transform of $h = \Psi(-\delta)$. Thus $H_e(a)$ is the slope response of the ETI system because

$$y(n) = x(n) \ominus (-h(-n)) \implies Y_e(a) = X_e(a) + H_e(a)$$

The \mathcal{A}_e transform has very similar properties with its counterpart \mathcal{A} used in DTI systems. The forward transform \mathcal{A}_e always yields a concave function $X_e(a)$. The ‘inverse’ transform $\mathcal{A}_e^{-1} : X_e(a) \mapsto \check{x}(n)$ with

$$\check{x}(n) \triangleq \bigvee_a X_e(a) + a(n) \leq x(n)$$

yields a convex signal that is equal to $x(n)$ if x is convex; otherwise $\check{x}(n)$ is the largest convex **lower envelope** of $x(n)$.

Let us return to (5) describing an ETI system and assume that $b_0 = 0$ and $b_m = \infty$ for $m > 0$. **1st-order system:** If $K = 1$, then the impulse and slope response are

$$h(n) = a_1 n - s(n) \xleftrightarrow{\mathcal{A}_e} H_e(a) = s(a_1 - a)$$

Thus this system acts as a ‘slope lowpass’ filter since it eliminates all linear trends in the input whose slope is $> a_1$.

For a **K^{th} -order system** with $K > 1$, applying \mathcal{A}_e to (5) yields

$$H_e(a) = \min[H_e(a) - a + a_1, \dots, H_e(a) - Ka + a_K, 0]$$

Thus $H_e(a) = s(a_0 - a)$ where $a_0 = \min_k \{a_k/k\}$, and the convex lower envelope of $h(n)$ is $\check{h}(n) = a_0 n - s(n)$. Thus a recursive K^{th} -order ETI system behaves, over long time scales, effectively as a 1st-order system.

4. ENVELOPE ESTIMATION

Here we develop an application of recursive DTI (ETI) systems where they can find upper (lower) envelopes of signals; e.g., see Figs. 2(d,e,f) and Fig. 3. Consider the problem of envelope detection in AM signals

$$x_{AM}(t) = [1 + \lambda \cos(\omega_a t)] \cos(\omega_c t), \quad \omega_a \ll \omega_c$$

Fig. 3a shows one period of a sampled AM signal $x(n) = x_{AM}(nT)$, where T is sampling period, with $\omega_a T = \pi/50$, carrier $\omega_c T = \pi/5$, and modulation index $\lambda = 0.5$. Consider the DTI system described by $y(n) = \max[y(n-1) + a_1, x(n)]$ with $a_1 < 0$. The output $y(n)$ is constrained to be $\geq x(n)$ and hence provides a type of upper envelope of $x(n)$. Computing $y(n)$ in forward time (see Fig. 3b), we see that downhill in between the consecutive peaks of the carrier $\cos(\omega_c t)$ $y(n)$ falls linearly with slope a_1 . Uphill $y(n)$ continues to fall between peaks, whereas it should rise. Hence, we also pass $x(n)$ through the anti-causal system $z(n) = \max[z(n+1) + a_1, x(n)]$ run backwards in time (see Fig. 3c). Then the final estimated upper envelope in Fig. 3d is the $\max y(n) \vee z(n)$ of the two outputs. Similarly, the ETI system $y(n) = \max[y(n) - a_1, x(n)]$ can yield a lower envelope of $x(n)$ by computing the recursive equation forward and backward in time and taking the min of the two outputs (see Figs. 3b,c,d). To maximize the smoothness of the resulting envelope we selected the slope parameter a_1 to match the average slope of the true envelope $f(t) = [1 + \lambda \cos(\omega_a t)]$ within time intervals equal to the carrier period $2\pi/\omega_c$. To avoid dependency on the location of such time intervals we also averaged over one half the period of $f(t)$ where $df/dt \leq 0$. This yielded

$$a_1 = -\frac{2\lambda\omega_c T \sin[2\pi(\omega_a/\omega_c)]}{\pi^2}$$

which has the value $a_1 = -0.0374$ for the example of Fig. 3. For envelope signals more general than a cosine, the same formula could be applicable if ω_a is the bandwidth of the envelope. It remains to be seen whether the above choice of a_1 is optimal according to some criterion.

The efficiency of 1st-order recursive DTI or ETI systems to estimate signal envelopes and their extremely small complexity (2 additions and 3 comparisons per output sample) makes them promising for AM detection and other applications of envelope detection.

5. EXTENSIONS TO 2-D SYSTEMS

2-D DTI systems $\Psi : x(n, m) \mapsto y(n, m)$ have as eigen-functions the signals $x(n, m) = an + bm + c$ because the corresponding outputs are $y(n, m) = an + bm + c + G(a, b)$ where the slope response

$$G(a, b) \triangleq \bigvee_{n, m} g(n, m) - an - bm, \quad (a, b) \in \mathbb{R}^2$$

is the 2-D \mathcal{A} transform of the impulse response $g(n, m) = \Psi[\delta(n, m)]$, and $\delta(n, m) = \delta(n) + \delta(m)$

is the 2-D zero impulse. Note that G has two slope variables: horizontal and vertical. The 2-D version of the transform \mathcal{A}^{-1} yields the signal

$$\hat{g}(n, m) \triangleq \bigwedge_a \bigwedge_b G(a, b) + an + bm \geq g(n, m)$$

which is the smallest concave upper envelope of g .

Causal DTI systems described by the equation

$$y(n, m) = x(n, m) \vee \bigvee_{\substack{k, \ell = 0 \\ k\ell \neq 0}}^K y(n-k, m-\ell) + a_{k\ell}$$

have an infinite 2-D impulse response $g(n, m)$ which is $-\infty$ if $n < 0$ or $m < 0$. As a 1st-order example consider the case $K = 1$ with $a_{10} = a_1$, $a_{01} = b_1$, and $a_{11} = -\infty$. Then the impulse response is

$$g(n, m) = a_1 n + b_1 m + s(n, m)$$

and the slope response is

$$G(a, b) = -s(a - a_1, b - b_1)$$

The min version of this 1st-order recursive equation (i.e., the ETI system) was used to compute the distance transform of binary images with respect to the city-block metric [5].

6. REFERENCES

- [1] G. Borgefors, "Distance Transformations in Arbitrary Dimensions", *Comp. Vision Graph. Image Process.*, 27, pp.321-345, 1984.
- [2] B. Lay, "Recursive Algorithms in Mathematical Morphology", *Acta Stereol.*, 6/III, pp.691-696, 1987.
- [3] P. Maragos and R. W. Schafer, "Morphological Filters - Parts I & II", *IEEE Trans. ASSP.*, 35:1153-1184, Aug. 1987.
- [4] T.A. Nodels & N.C. Gallagher, "Median Filters: Some Modifications and Their Properties", *IEEE Trans. ASSP.*, 30:739-746, Oct. 1982.
- [5] A. Rosenfeld and J. L. Pfaltz, "Sequential Operations in Digital Picture Processing", *J. ACM*, 13, pp. 471-494, Oct. 1966.
- [6] J. Serra, *Image Analysis and Mathematical Morphology*, NY: Acad. Press, 1982.
- [7] S. R. Sternberg, "Grayscale Morphology," *CVGIP*, 35:333-355, 1986.

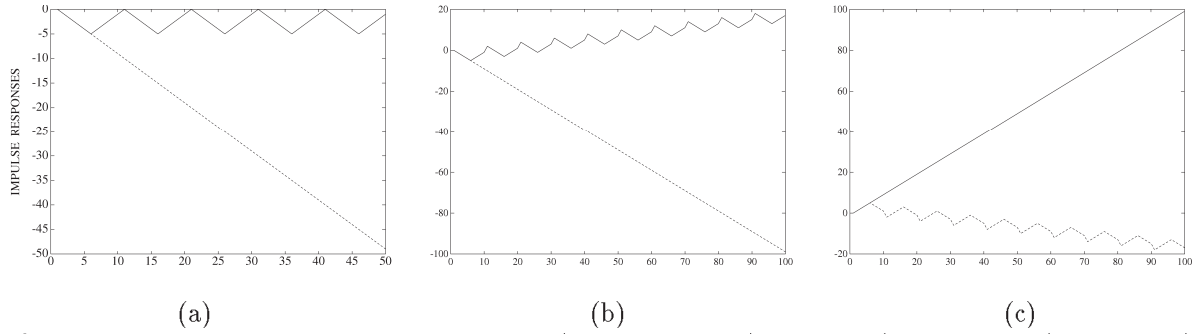


Figure 2: Impulse responses of recursive DTI (solid upper line) and ETI (dot lower line) systems. (a) $a_k = |5 - k| - 5$, $k = 1, \dots, 10$. (b) a_k as in (b) except $a_{10} = 2$. (c) a_k are negatives of the ones in (b).

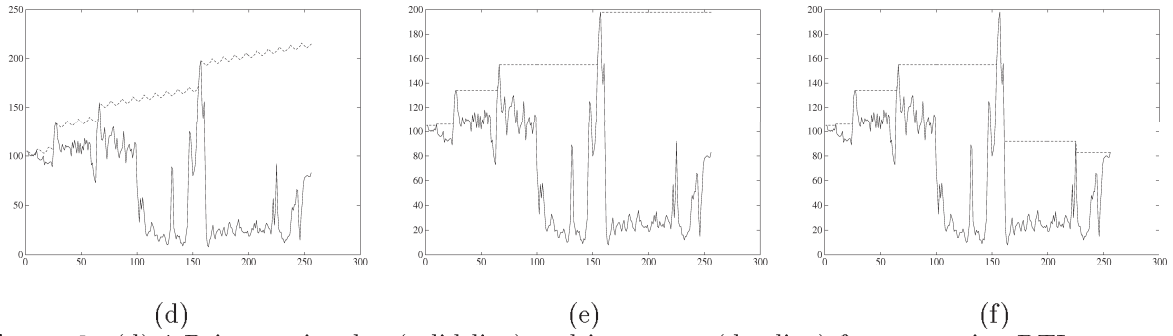


Figure 2: (d) 1-D image signal x (solid line) and its output (dot line) from recursive DTI system of Fig. 2(b). (e) Image x and its output from system $y(n) = \max[y(n-1), x(n)]$ in forward time. (f) Min of outputs of system in (e) computed in forward and backward time.

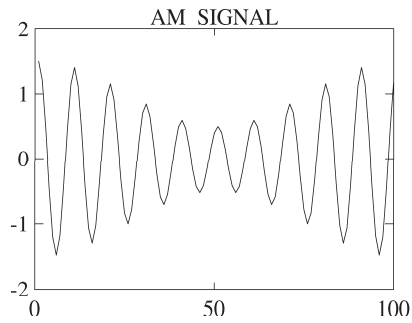


Figure 3(a)

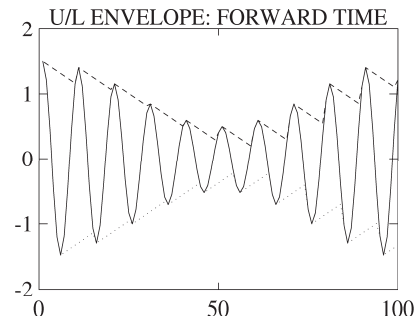


Figure 3(b)

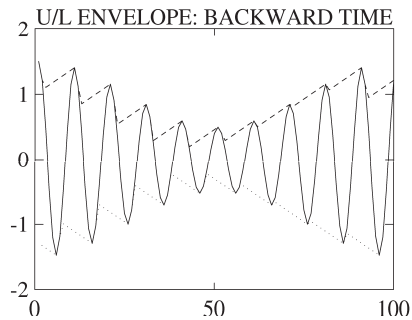


Figure 3(c)

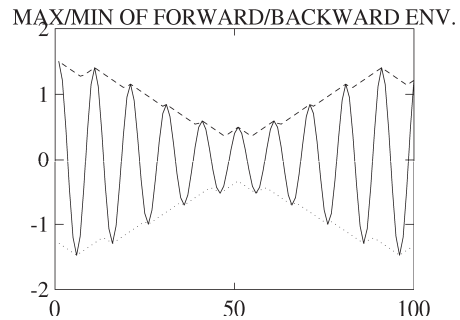


Figure 3(d)

Figure 3: Envelope estimation in AM signals via 1st-order recursive DTI and ETI systems.