Waveform Recognition in the Presence of Domain and Amplitude Noise

Mohamad A. Akra and Sanjoy K. Mitter, *Fellow, IEEE*

***Abstract*—In this paper, we discuss the problem of recognizing single-dimensional, real-valued, functions in the presence of domain noise (i.e., noise that affects the domain rather than the amplitude). This problem is inspired by the field of on-line character recognition where it is more natural to view the hand as deforming the domain of the character rather than adding noise to its amplitude. The results obtained illutrate the difficulties one faces when dealing with both domain and amplitude deformation of waveforms or images. Our major result is a set of sufficient conditions that a recognition metric has to satisfy. Examples of metrics that satisfy these conditions, and hence are appropriate for recognition when the deformation affects the domain rather than the amplitude, include the supnorm metric and the total variation metric. Furthermore, we extend the results to the case when a waveform is corrupted by both amplitude and domain deformation.**

***Index Terms*—Character recognition, metric, domain, amplitude, deformation.**

I. Mathematical Preliminaries

For the sake of our analysis, we assume that the functions under consideration have bounded support. Alternatively, and without loss of generality, we can assume these functions to be defined on the unit interval. Let H be the space of real-valued functions of bounded variation, defined over the unit interval I, satisfying the following.[1]

1) At every point t in the interval $[0, 1]$, the left- and right-hand limits exist.
2) $\forall t \in (0,1), \forall f \in H, f(t) = f(t+0+)$ or $f(t) = f(t-0)$.
3) The values taken by f at 0 and 1 are arbitrary.

Let $h_1, h_2, \cdots, h_M$ be elements of H called *hypotheses*. Let $g \in H$ be the *received* signal. Determine a plausible detection rule to determine the most likely h_i that was transmitted.

The traditional approach in signal transmission theory is to assume that g was obtained from one of the h_i's through amplitude deformation (additive or multiplicative). In other words, we make M hypotheses $H_1, H_2, \cdots, H_M$, where under hypothesis H_i (assuming additive noise)

$$g(t) = h_i(t) + n(t), \qquad 0 \le t \le 1$$

Manuscript received January 23, 1995; revised April 15, 1996.

M. A. Akra was with the Department of Electrical and Computer Engineering, American University of Beirut, Beirut, Lebanon. He is now with the Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

S. K. Mitter is with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

Publisher Item Identifier S 0018-9448(97)00113-2.

[1] These properties result in functions having finite ϵ-oscillations and a countable number of discontinuities.

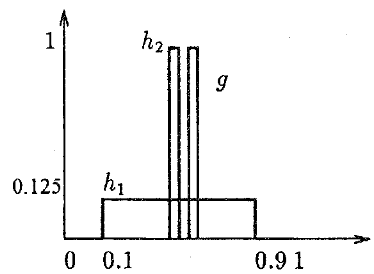


Fig. 1. Counter example showing weakness of L_2 metric.

and $n(t)$ denotes the amplitude noise, lying in some suitable space. Next, we define a cost functional for $n(t)$. Finally, we select the hypothesis that minimizes the cost.

If the L^2 metric were used as a cost on H, for instance, the detection rule would be to choose the h_i that minimizes

$$\int (g(t) - h_i(t))^2 \, dt.$$

To use a probabilistic interpretation, this is the rule that corresponds to maximum-likelihood detection when $n(t)$ is modeled as Gaussian white noise.

Unfortunately, comparing signals based on their amplitude difference only does not satisfy our intuition. Consider, for instance, the following example:

$$h_1(t) = \begin{cases} 0.125, & 0.1 \le t \le 0.99 \\ 0, & \text{otherwise} \end{cases}$$

$$h_2(t) = \begin{cases} 1, & 0.5 \le t \le 0.51 \\ 0, & \text{otherwise} \end{cases}$$

$$g(t) = \begin{cases} 1, & 0.52 \le t \le 0.53 \\ 0, & \text{otherwise.} \end{cases}$$

In this example, g is closer to h_1 in the L^2-sense, whereas it is closer to h_2 in any intuitive sense (see Fig. 1).

To explain the apparent inconsistency, note that the L^2 metric assigns a large cost to domain deformation. On the other hand, it seems that humans do not assign such a cost to domain deformation when comparing objects. So it seems that we should formulate the problem in a way that accounts for domain deformation in addition to amplitude deformation.

First, we focus on domain deformation only. Then, we combine both domain and amplitude deformations. A naive attempt to solve the problem is to say: under hypothesis H_i we have

$$g(t) = h_i(t + w(t)) \tag{1}$$

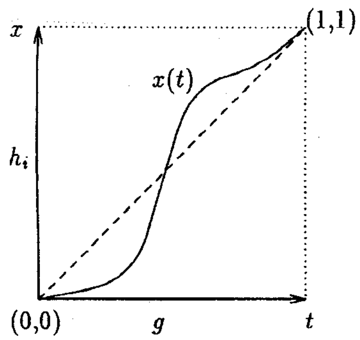


Fig. 2. Domain deformation shown as a curve in the product domain of g and h_i.

and then use the cost function

$$\int w(t)^2\,dt$$

as a measure of similarity (in the intuitive sense). However, this function has annoying asymmetry. The reason is, if we write $g(t) = h_i(t + w_i(t))$, and alternatively, $h_i(t) = g(t + w_i'(t))$, then it is not true that

$$\int w_i(t)^2\,dt = \int w_i'(t)^2\,dt.$$

Furthermore, if we decide later to shift our interest to functions having infinite support, it is possible for a small domain shift to cause a large increase in the cost function (consider $\cos(t)$ and $\cos(t+\epsilon)$), which means it is not translation-invariant. A more convenient relation than (1) is

$$g(t) = h_i(x(t)) \quad \text{or} \quad h_i(t) = g(x^{-1}(t)) \tag{2}$$

assuming x is invertible. Using (2) we can immediately see that for a cost functional to be symmetric, it has to remain the same when x is replaced by x^{-1}. These ideas will be made more concrete in the following section.

II. Domain Deformation: Theory

Assume that under hypothesis H_i

$$g(t) = h_i(x(t)) = (h_i \circ x)(t)$$

where $x(t)$ is an order-preserving homeomorphism of the unit interval I onto itself. We will show later that a solution x for the above equation exists if and only if g and h_i have the same sequence of extrema.

The reader should note that, throughout the coming discussion, we will be dealing with three different spaces: X, H, and W. First, we will define X, and then we will present a few lemmas to gain some understanding about the space X.

Let X be the space of all order-preserving homeomorphisms of the unit interval I onto itself. Then, the following two lemmas are well known.

Lemma 2.1: A function x is an element of X if and only if it can be represented as a continuous, strictly increasing function joining the origin to the point $(1,1)$ of $I \times I$ (see Fig. 2). The horizontal axis in Fig. 2 (also called the t-axis) is the domain of g, and the vertical axis (also called the x-axis) is the domain of h_i.

Note that the functions g and h_i are to be visualized normal to the plane of Fig. 2.

Lemma 2.2: The pair $(X, \circ)$ is a group.

Note that the inverse x^{-1} is a reflection of x around the diagonal of $I \times I$.

Lemma 2.3: The space X, viewed as a set, is convex.

Proof: By convex we mean, if x_1 and x_2 belong to X, then so do $\alpha x_1 + (1-\alpha)x_2$ for all α in $[0,1]$. Let $x = \alpha x_1 + (1-\alpha)x_2$. Then x is continuous, being a weighted sum of continuous functions. Letting $t_1 < t_2$, we get

$$\begin{aligned} x(t_2) &= \alpha x_1(t_2) + (1-\alpha)x_2(t_2) \\ &> \alpha x_1(t_1) + (1-\alpha)x_2(t_1) \\ &= x(t_1). \end{aligned}$$

Finally, $x(0) = 0$, and $x(1) = 1$. Hence, x is in X.

Next, we define the second space of interest, H, and its subspace H_g.

Definition 2.4: Let H be the space of real functions of bounded variations defined over I. Let g be an element of H. H_g is defined to be the set of all functions in H which can be obtained from g through order-preserving, homeomorphic, domain deformation. In other words

$$H_g = \{f \colon g = f \circ x, x \in X\}.$$

Note that we already have an onto mapping from X to H_g, mapping x to $g \circ x^{-1}$. However, this mapping is not one-to-one in general. Consider, for instance, the constant function $g(t) = 1$. Then, H_g is the singleton $\{g\}$, since $1 \circ x^{-1} = 1$ for all x in X. By removing from X the "redundant" x's, the onto mapping becomes a bijection. For the case $g = 1$, the redundant deformations are all the deformations except the identity deformation $x = i$. The following lemmas characterize the redundant domain deformations for a general function g.

Lemma 2.5: Let f and g be both strictly increasing (decreasing) functions. Assume that the relation $g = f \circ x$ has a solution in X. Then, that solution is unique.

Proof: Assume the contrary. Let x_1 and x_2 be two different solutions of the equation $g = f \circ x$. Then, there exists a point a at which they differ. Let $x_1(a) = b_1$, $x_2(a) = b_2$. However, $g(a) = f[x_1(a)] = f[x_2(a)]$. Therefore, $f(b_1) = f(b_2)$ for $b_1 \neq b_2$. This is a contradiction, since f is strictly monotonic. Hence, the solution x is unique.

Lemma 2.6: Let $f, g \in H$. Let $g = f \circ x$ have a solution in X. Then, x is a bijection between the local maxima (minima) of g and the local maxima (minima) of f.

Proof: Let a be a point at which g has a local maximum. Then, there is an open neighborhood of a, $B(a) \subset I$, such that $g(t) \leq g(a)$ for all t in $B(a)$.[2] Consider $x[B(a)]$. It is open since x is a homeomorphism, and contains $x(a)$. For all u in $x[B(a)]$, we have $f(u) \leq f[x(a)]$, otherwise $x^{-1}(u)$ would be a point in $B(a)$ such that $g[x^{-1}(u)] \geq g(a)$. Hence, $x(a)$ is a local maximum of f. Similarly, if b is a local maximum of f, we can prove that $x^{-1}(b)$ is a local maximum of g. Finally, the minima can be treated in an analogous way. Note that the

[2] To handle the extreme points 0 and 1, view I as a subspace of R, where the topology is inherited by taking the intersections of open sets of R with I.

theorem holds for straddle points as well. However, the proof for this case is only a small modification of the one above.

Theorem 2.7: Suppose that g is not constant on any subinterval of I, and that the relation $g = f \circ x$ has a solution in X. Then, the solution is unique on the whole interval.

Proof: Let $a_1 < a_2 < \cdots$ be the points at which g has extrema. Let $b_1 < b_2 < \cdots$ be the points at which f has extrema. Let $g = f \circ x$ has a solution. Then x in Fig. 2 should pass through the points $(0,0), (a_1, b_1), (a_2, b_2), \cdots, (1,1)$, otherwise x would not be increasing. Furthermore, between any two consecutive points $(a_i, b_i), (a_{i+1}, b_{i+1})$, f and g are either both increasing or decreasing. As a consequence of Lemma 2.5, the solution is unique in $[a_i, a_{i+1}]$. Therefore, the solution is unique on the whole interval.

Corollary 2.8: Suppose that g is not constant on any subinterval of I. Then there exists a bijection between H_g and X, mapping f in H_g to the unique x in X, satisfying $g = f \circ x$.

Corollary 2.9: Let g be constant on the ordered intervals $(c_1, c_2), (c_3, c_4), \cdots$. Let f be constant on the ordered intervals $(d_1, d_2), (d_3, d_4), \cdots$. Then

1) any valid solution of $f = g \circ x$ maps (c_i, c_{i+1}) to (d_i, d_{i+1}) for all i;
2) outside these subintervals, x is uniquely determined.

Throughout the work, we will choose the solution x that maps (c_i, c_{i+1}) to (d_i, d_{i+1}) *linearly*. As a consequence of this, the solution to $g = f \circ x$ becomes unique. We denote by X_g the subset of X obtained after deleting every x which is not linear in the constant subintervals of g.

Our goal has been to define a distinction function d^* over $H_g \times H_g$, which handles domain deformation in a way analogous to the way L^2 or other metrics handle amplitude deformation. The direct consequence is that d^* becomes a metric over H. The second consequence is that $d^*(f, f \circ x)$ becomes a function of x only.

The second consequence implies that $d^*(f_1, f_2)$ should not change if we deform both domains by the same deformation (invariance to composition). The equivalent assumption in the classical case is to say the amplitude noise cost is dependent on the difference $g(t) - h_i(t)$ only (invariance to translation).

Let us denote $d^*(f, f \circ x)$ by $\langle x \rangle$. Then, the properties of a distinction function are

1) $x \in X$ is the identity function if and only if $\langle x \rangle = 0$;
2) if $x \in X$, then $\langle x \rangle = \langle x^{-1} \rangle$;
3) if x_1 and x_2 are in X, then

$$\langle x_1 \rangle + \langle x_2 \rangle \geq \langle x_1 \circ x_2 \rangle.$$

Clearly, if $\langle x \rangle = d^*(f, f \circ x)$ is a distinction function, then so is $\lambda \langle x \rangle$ for any positive constant λ.

Note that the distinction function d^* defined on H_g induces a metric on X_g. In other words, if x_1 and x_2 are in X_g, then $d_{X_g}(x_1, x_2) = \langle x_1 \circ x_2^{-1} \rangle$ is a metric on X_g. Furthermore, the two metric spaces H_g and X_g become very similar in the sense of the following theorem:

Theorem 2.10: Let d^* be a function over $H_g \times H_g$. Let d be a function over $X_g \times X_g$ defined as follows:

$$d(x_1, x_2) = d^*(f_1, f_2) \quad \text{such that } g = f_1 \circ x_1 = f_2 \circ x_2.$$

Then, d^* is a distinction function if and only if d is a composition-invariant metric. Furthermore, if d is a composition invariant metric on X, the mapping

$$\begin{aligned} b\colon (H_g, d^*) &\to (X, d) \\ f &\to x \in X_g\colon g = f \circ x \end{aligned}$$

is an *isometric embedding* of H_g in X.

Proof: The mapping b is well defined since x is uniquely determined (see Corollary 2.9 and the following remark). Furthermore

$$d^*_{H_g}(f_1, f_2) = d_X(b(f_1), b(f_2)).$$

Hence, b is an isometric embedding of H_g in X, or an isometry between H_g and X_g. Since b is a bijection between H_g and X_g, then d is a metric if and only if d^* is a metric. Furthermore, let $g = f_1 \circ x_1 = f_2 \circ x_2$, then $d^*(f_1, f_2) = d^*(g, g \circ x_2^{-1} \circ x_1)$ if and only if $d(x_1, x_2) = d(i, x_2^{-1} \circ x_1)$.

Isometry has important consequences since isometric spaces are identical in all respects except for the nature of their elements, which is inessential [2].

The question is: "Does there exist a valid distinction function?" Happily, the answer is "Yes!" Later in the discussion we will present several examples. For now, the reader can verify that the following is one such function:

$$d^*(f, f \circ x) = \sup_{t \in I} |x(t) - t|.$$

In what follows, we define the third space of interest W.

Definition 2.11: Let X be the space of order-preserving homeomorphisms defined earlier. Define W to be the noise part of the domain deformation. In other words

$$W = \{w(t)\colon w(t) = x(t) - t, x \in X\}.$$

The bijection between W and X is immediate, and the convexity of $(X, +)$ implies the convexity of $(W, +)$. However, $(W, \circ)$ is not a group. Also, $(W, +)$ is not a linear space, so we cannot define a norm over it. However, if we extend W by including all continuous functions w satisfying $w(0) = w(1) = 0$, then the extended space becomes a linear space. Any norm on the extended space may be inherited by W. Alternatively, any metric on the extended space can also be used to metrize W.

Theorem 2.12: Let d' be a metric over W, induced by defining a norm over the extended W. Let d be a function over $X \times X$ defined as follows:

$$d(x_1, x_2) = d'(x_1 - i, x_2 - i).$$

Let x_1 and x_2 be two points in X. Then, the following mapping:

$$\begin{aligned} \Gamma\colon [0,1] &\to X \\ \alpha &\mapsto \alpha x_1 + (1 - \alpha) x_2 \end{aligned}$$

is a shortest path joining x_1 to x_2 in (X, d). Furthermore, the mapping

$$\begin{aligned} b\colon X &\to W \\ x &\mapsto w = x - i \end{aligned}$$

is an isometry between (X, d) and (W, d').

Proof: b is a bijection that preserves distances, hence it is an isometry. Let d' be a norm in W. Consider the mapping

$$\begin{aligned}\Gamma'\colon [0,1] &\to W\\ \alpha &\mapsto \alpha w_1 + (1-\alpha)w_2.\end{aligned}$$

Then Γ' is a shortest path in W, since

$$\begin{aligned}&d'(w_1, \alpha w_1 + (1-\alpha)w_2) + d'(\alpha w_1 + (1-\alpha)w_2, w_2)\\ &\quad = \|(1-\alpha)w_1 - (1-\alpha)w_2\| + \|\alpha w_1 - \alpha w_2\|\\ &\quad = (1-\alpha)\|w_1 - w_2\| + \alpha\|w_1 - w_2\|\\ &\quad = \|w_1 - w_2\|\\ &\quad = d'(w_1, w_2).\end{aligned}$$

Since (X, d) and (W, d') are isometric, shortest paths in one space are also shortest paths in the other. Therefore, $\alpha x_1 + (1-\alpha)x_2$ is a shortest path in (X, d), joining x_1 to x_2. Note that the converse is not true. In other words, $\alpha x_1 + (1-\alpha)x_2$ can be a shortest path in a metric space (X, d) without d being a norm. An example of this is $X = I$ with the metric

$$d(x_1, x_2) = \left|\int_{x_1}^{x_2} \sqrt{1+4t^2}\, dt\right|.$$

To summarize what has been done so far, consider the three spaces:

1) H_g, the space of observations, when only domain deformation is involved;
2) X_g, the space of domain deformations, constructed in a way to make a bijection with H_g;
3) W_g, the space of domain noise.

Then the following theorem illustrates the links between the various spaces.

Theorem 2.13: Let d' be a norm metric over W, i.e., $d'(w_1, w_2) = \|w_1 - w_2\|$. Let d be the corresponding metric over X, i.e., $d(x_1, x_2) = d'(x_1 - i, x_2 - i)$. Assume further that d is composition-invariant, i.e., $d(x_1, x_2) = d(i, x_1 \circ x_2^{-1})$. Let d^* be the corresponding metric over H_g, i.e.,

$$d^*(f_1, f_2) = d(x_1, x_2)\colon g = f_1 \circ x_1 = f_2 \circ x_2.$$

Then, d^* is a distinction function over H_g. Furthermore, the following mapping

$$\begin{aligned}\Gamma\colon [a,b] &\to H_g\\ \alpha &\mapsto \Gamma(\alpha) = f\colon g = f \circ (\alpha x_1 + (1-\alpha)x_2)\end{aligned}$$

is a shortest path in (H_g, d^*) joining f_1 to f_2.

Let us list some examples of distinction functions over H_g:

Supremum:

$$d^*(f_1, f_1 \circ x) = \langle x\rangle = \sup_t |x(t) - t|$$

Maxmax:

$$d^*(f_1, f_1 \circ x) = \max_t(x(t) - t) + \max_t(t - x(t))$$

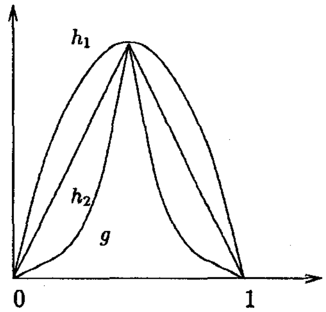


Fig. 3. Binary hypothesis of Example 1.

Variation:

$$d^*(f_1, f_1 \circ x) = \text{total variation of } (x - t) = V(x - t).$$

Note that if x is differentiable,

$$V(x - t) = \int_t |\dot{x} - 1|\, dt$$

III. Domain Deformation: Application

Given M functions $h_1, h_2, \cdots, h_M$ in H_g and a specific choice of the distinction function, we define the recognition rule as follows: Choose h_i that minimizes $\langle x_i\rangle$, where $g = h_i \circ x_i$.

Example 1:

$$\begin{aligned}h_1(t) &= \sin(\pi t), && 0.0 \le t \le 1.0\\ h_2(t) &= \begin{cases} 2t, & 0.0 \le t \le 0.5\\ 2(1-t), & 0.5 \le t \le 1.0\end{cases}\\ g(t) &= \begin{cases} 4t^2, & 0.0 \le t \le 0.5\\ 4(1-t)^2, & 0.5 \le t \le 1.0.\end{cases}\end{aligned}$$

Decide whether g is closer to h_1 or to h_2 in all of the above metrics (see Fig. 3).

Solution 1: Let $g = h_1 \circ x_1$. Solving, we get $x_1(t) = \sin^{-1}(4t^2)/\pi$ for $0 < t < 1/2$ with symmetry w.r.t. $(1/2, 1/2)$. To find the *supremum* distance we set $\dot{x}_1 = 1$ to obtain (using Maple)

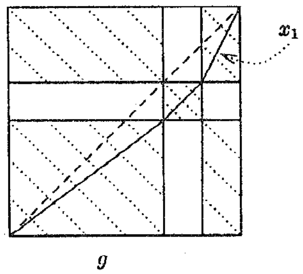
$$\langle x_1\rangle = 0.187.$$

Let $g = h_2 \circ x_2$. Solving, we get $x_2(t) = 2t^2$ for $0 < t < 1/2$ with symmetry w.r.t. $(1/2, 1/2)$. Similarly, we set $\dot{x}_2 = 1$ and we obtain the *supremum* distance

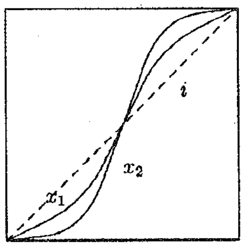
$$\langle x_2\rangle = 0.125.$$

Since $\langle x_1\rangle > \langle x_2\rangle$, we say that g resembles h_2 more than h_1 in the *supremum* metric. The *maxmax* metric gives the same answer in this particular case since $\max(x(t) - t) = \max(t - x(t))$ (see Fig. 4). Finally, the *total variation* metric also gives the same result since it is equal to twice the *maxmax* metric in this case.

In practice, one can view the various solutions of the equation $g(t) = h_i(x(t))$ by drawing the contour plot of the surface $|g(t) - h_i(x)|$ in the plane of Fig. 2. The kernel of the surface (i.e., the regions where it is zero) contains all the solutions. This kernel can be viewed using a mathematical



Kernel for $g - h_2$ from Figure 1



Domain deformations from Example 1

Fig. 4. Calculation of deformation.

package like MATLAB. Out of that kernel we choose the curve x that connects the origin $(0, 0)$ to the opposite corner $(1, 1)$ with minimum $\langle x \rangle$. Note that unless g is constant for some interval $(a, b) \subset I$, the kernel will be a finite collection of curves. However, only one of them will connect the origin to the opposite corner. If g is constant for some interval, the corresponding part of the kernel will be a rectangle (see Fig. 4). In all of the defined metrics, the reader can verify that taking the diagonal of that rectangle only and dropping the rest results in the minimum $\langle x \rangle$, which justifies the criteria made earlier in this section.

IV. Conclusion

In the classical amplitude-only deformation case, the invariance of the metric with respect to the order of the incremental noise addition was guaranteed by choosing a metric which is a function of the sum of the incremental noises, or, alternatively, a function of the difference between the transmitted and the received signal. This choice guarantees invariance because, obviously, the sum is invariant with respect to the order of the elements being added. Similarly, in the domain-only deformation case, the invariance was guaranteed by choosing a metric that depends on the overall deformation $x(t)$, regardless of the individual domain deformations. Also, by forcing an isometric embedding with the space W, we were able to guarantee that the shortest path of differential deformations dx_i, is actually the line segment connecting the transmitted and received signals in the space W.

When considering both amplitude and domain deformations, special care is needed to make sure that, when the two types of deformation intermix, the cost function is not affected. This will be the subject of the following discussion.

V. Amplitude and Domain Deformation: Theory

In the following theorem, which is easy to verify, we make a slight generalization:

Theorem 5.1: Let H be the space of all functions of bounded variations whose domain is I. Let X be the space of homeomorphisms over I. Let d be a function over $H \times H$ defined as follows:

$$d(f_1, f_2) = \langle x \rangle \text{ if there exists } x \text{ in } X \text{ such that}$$
$$f_1 = f_2 \circ x$$
$$d(f_1, f_2) = \infty, \quad \text{otherwise.}$$

Then, d is a generalized metric over H. Furthermore, we call d a domain metric.

One can visualize H as a collection of disjoint slices where the distance d between any two slices is ∞. In fact, these slices are called *path components* in the topological sense, since any pair of points belonging to the same slice are path-connected, and no two points belonging to different slices can be connected.

On the other hand, we can define over H an amplitude metric as follows:

Definition 5.2: Let d be a metric over H such that for any pair of points f_1 and f_2 in $H, d(f_1, f_2)$ is a function of $f_1 - f_2$ only. Then, d is called an amplitude metric.

Note that d is a pseudonorm but not necessarily a norm (e.g., consider the discrete metric). The most widely used examples of amplitude metrics are the so-called L^p metrics, where for $p \geq 1$:

$$d_2(f_1, f_2) = \left(\int_t |f_1(t) - f_2(t)|^p \, dt \right)^{1/p}.$$

Having defined suitable metrics for domain-only and amplitude-only deformations, how can we use these results to develop a suitable metric for mixed amplitude and domain deformations?

One way to do this is to note that the metrics obtained are also path metrics. In other words, they are defined over path-connected spaces in such a way that $d(p, q)$ is equal to the length of the shortest path connecting p to q [3]. Now envision the space H with two metrics: d_1, an amplitude metric, and d_2, a domain metric. We can view H with these two metrics as a city with two modes of transportation, e.g., *walking* and *train*. $d_1(A, B)$ can be thought of as the time needed to walk from A to B, while $d_2(A, B)$ is the time needed to travel by train. Note that unless A and B are train stations for the same train line, $d_2(A, B) = \infty$. When both means of transportation are possible, we can define the travel time between A and B as that time spent when the optimum use of both means is made. Along the same lines, we can define the cost of deforming a function f_1 to another function f_2 as the cost of optimum combination of amplitude and domain deformations. The recognition problem becomes that of finding the hypothesis that requires the least such cost. In what follows, we present the above analysis in a more precise way.

Let Γ: $[a,b] \to (H,d_1)$ be a path in (H,d_1) joining f_1 to f_2. For any partition of $[a,b]$ given by

$$P = \{t_0, t_1, \cdots, t_m\}$$

the points $\Gamma(t_0), \Gamma(t_1), \cdots, \Gamma(t_m)$ are the vertices of an inscribed polygon. The length of this polygon is denoted $\Delta_\Gamma(P)$ and is defined to be the sum

$$\Delta_\Gamma(P) = \sum_{k=1}^{m} \min(d_1(\Gamma(t_{k-1}), \Gamma(t_k)), d_2(\Gamma(t_{k-1}), \Gamma(t_k))).$$

Definition 5.3: The distance between f_1 and f_2 is the least upper bound of $\Delta_\Gamma(P)$ over all partitions P of $[a,b]$, i.e.,

$$d(f_1, f_2) = \sup\{\Delta_\Gamma(P) \colon P \in \mathcal{P}[a,b]\}.$$

Theorem 5.4: The distance function of Definition 5.3 is a metric.

The fact that d is a metric comes from its very definition. The reader can refer to Shreider [5] for a discussion on this approach for defining distances. Nevertheless, we will include the proof here for completeness.

Proof:

1) (*Identity*) If $f_1 = f_2$, then the constant function $\Gamma(\alpha) = f_1$ is continuous and defines a path of zero length. Therefore, $d(f_1, f_2) = 0$. To prove the converse, assume that $d(f_1, f_2) = 0$. Then, for any partition of the shortest path, we have to have either

$$d_1(\Gamma(t_{k-1}), \Gamma(t_k)) = 0$$

or

$$d_2(\Gamma(t_{k-1}), \Gamma(t_k)) = 0.$$

However, if one of the distances is zero the two points are coincident, hence the other distance should be zero as well. Therefore, all points of the partition are the same. Consequently, $f_1 = f_2$.

2) (*Symmetry*) follows from the symmetry of d_1 and d_2.

3) (*Triangle Inequality*) This is the easiest to prove. For if there exists a point f such that

$$d(f_1, f) + d(f, f_2) < d(f_1, f_2)$$

it means we can obtain a shorter path joining f_1 to f_2 by having the path pass through f, a contradiction.

The question becomes, given f_1 and f_2, how do we find the shortest path between them? This is not an easy question. Assume that the shortest path is an alternate sequence of domain and amplitude deformations such that, starting from f_1 we have:

$$f_1 \overset{n_1}{\to} f_1 + n_1 \overset{x_1}{\to} (f_1 + n_1) \circ x_1 \overset{n_2}{\to} (f_1 + n_1) \circ x_1 + n_2 \cdots \to f_2$$

where all n_i's, except possibly n_1, are different from 0. Similarly, all x_i's are different from the identity. Then we can describe any path emanating from f_1 in the following recursive form:

$$p_1 = f_1 + n_1,$$
$$p_{i+1} = p_i \circ x_i + n_i.$$

The path length would be

$$\Delta_\Gamma = \sum_i \langle x_i \rangle + \|n_i\|.$$

Our problem becomes: Find x_i and n_i, for $i = 1, 2, \cdots$ that minimize

$$\sum_i \langle x_i \rangle + \|n_i\|$$

subject to

$$\lim_{i \to \infty} p_i = f_2.$$

However, the problem can be simplified remarkably if we choose the amplitude norm $\| \;\|$ also to be composition-invariant. In that case, we can prove the following theorem:

Theorem 5.5: Let f_1 and f_2 be two functions defined over the unit interval I. Assume that f_2 was obtained from f_1 by an alternate sequence of domain and amplitude deformations, $\{x_i\}$ and $\{n_i\}$. Assume that the path length corresponding to a domain deformation is $\langle x_i \rangle$ and that corresponding to an amplitude deformation is $\|n_i\|$. Let l be the length of the shortest path joining f_1 to f_2. Assume that the amplitude norm $\| \;\|$ is composition-invariant. Then l does not increase if we restrict the number of deformations of each kind to be at most one. In other words, we can write

$$f_2 = f_1 \circ x + n$$
$$l = \min_x(\langle x \rangle + \|f_2 - f_1 \circ x\|)$$

Finally, $d(f_1, f_2) = l$ defines a metric.

To prove it, however, we need one lemma.

Lemma 5.6: Let H be the space of functions defined over I. Let f_1 and f_2 be two elements of H. Let d_1 be a composition-invariant amplitude metric over H. Let d_2 be a domain metric over H. Let (x_1, n_1) be the shortest pair of domain-amplitude deformations taking f_1 to f_2, i.e., if we let

$$D_1 = \{(x,n) \colon f_2 = f_1 \circ x + n\}$$

then

$$\langle x_1 \rangle + \|n_1\| \le \langle x \rangle + \|n\| \quad \forall (x,n) \in D_1.$$

Similarly, let (n_2, x_2) be the shortest pair of amplitude-domain deformations taking f_1 to f_2, i.e., if we let

$$D_2 = \{(n,x) \colon f_2 = (f_1 + n) \circ x\}$$

then

$$\|n_2\| + \langle x_2 \rangle \le \|n\| + \langle x \rangle \quad \forall (n,x) \in D_2.$$

Hence

$$\|n_1\| + \langle x_1 \rangle = \|n_2\| + \langle x_2 \rangle \quad \forall (n,x) \in D_2.$$

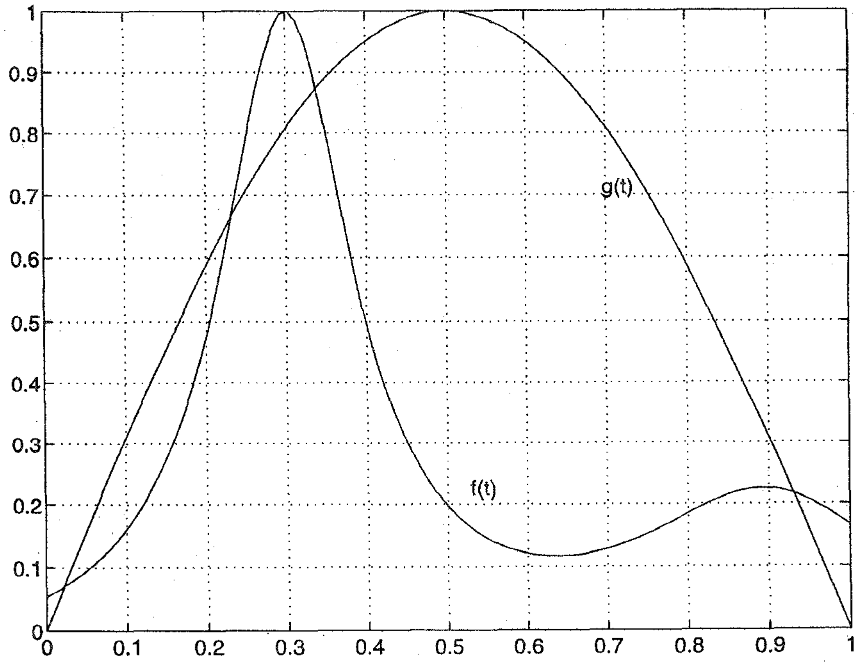


Fig. 5. $f(t) = 1/(x-3)^2 + 0.01 + 1/(x-9)^2 + 0.04 - 6$ and $g(t) = \sin(\pi t)$.

Proof: From the definition of x_1 and x_2 we can write

$$x_1 = \arg\min_x \langle x \rangle + \|f_2 - f_1 \circ x\|$$
$$x_2 = \arg\min_x \langle x \rangle + \|f_2 \circ x^{-1} - f_1\|.$$

However, the amplitude metric $\|\cdot\|$ is composition-invariant. Hence

$$\|f_2 \circ x^{-1} - f_1\| = \|(f_2 \circ x^{-1} - f_1) \circ x\| = \|f_2 - f_1 \circ x\|.$$

Therefore

$$\min_x \langle x \rangle + \|f_2 - f_1 \circ x\| = \min_x \|f_2 \circ x^{-1} - f_1\|$$

i.e.,

$$\langle x_1 \rangle + \|n_1\| = \langle x_2 \rangle + \|n_2\|.$$

Note that n_1 need not be the same as n_2, only the norm is the same. Now, we are in a position to prove Theorem 5.5.

Proof: Let $\{x_i\}$ and $\{n_i\}$ be the sequence of deformations corresponding to the shortest path joining f_1 to f_2. Consider a particular domain-amplitude deformation pair (x_k, n_k), joining $\Gamma(t_{k-1})$ to $\Gamma(t_k)$. Then, using Lemma 5.6, we can replace (x_k, n_k) by another amplitude-domain deformation pair (n'_k, x'_k) without increasing the length. Finally, n'_k can be added to n_{k-1}, and x'_k can be composed with x_{k+1} thereby reducing the number of points on the partition by 1. Repeating the process, we end up with one pair of amplitude-domain, or domain-amplitude, deformations without increasing the length. Obtaining one pair is equivalent to obtaining the other (see Lemma 5.6). Since every partition of the shortest path results in a polygon having the same length as one pair of domain-amplitude deformations, so is the supremum of all partition lengths. Q.E.D..

VI. Domain and Amplitude Deformation: Application

To illustrate the consequences of Theorem 5.5, assume that we are using the supremum metric for both amplitude and domain deformations. Given two functions f_1 and f_2, Theorem 5.5 implies that the following is also a metric over H

$$d(f_1, f_2) = \inf_x \left(\lambda_1 \sup_t |x(t) - t| + \lambda_2 \sup_t |f_2(t) - f_1(x(t))|\right) \tag{3}$$

where λ_1 and λ_2 are positive weights included for convenience.

In the literature on probability metrics, the above discovered metric[3] is called the Skorokhod metric [4]. Using it to solve the recognition problem of Fig. 1, we find after some thought that

$$d(g, h_1) = 0.875\lambda_2,$$
$$d(g, h_2) = \min(0.02\lambda_1, \lambda_2).$$

which implies that g is closer to h_2 (in agreement with the human sense) unless $\lambda_1/\lambda_2 > 43.75$. If the last inequality is true, it means that we are assuming a huge cost for domain deformation, and in that case only it is reasonable to say that g is closer to h_1.

Note that we have used the *supremum* norm in (3), since it is easier to deal with analytically. However, when running computer minimization algorithms that use gradient-descent-type techniques, it is more convenient to use the *total variation* norm instead.

When the functions of interest cannot be expressed in analytic forms, we can discretize them as follows, and then use a software package such as MATLAB to find the solution.

[3] It is important to note that H is not complete under this metric. However, there is an equivalent metric which makes H complete. We can prove that if $f_n \in H$ converges in the Skorokhod metric to a continuous function f then $\sup_t |f_n(t) - f(t)| \to 0$.

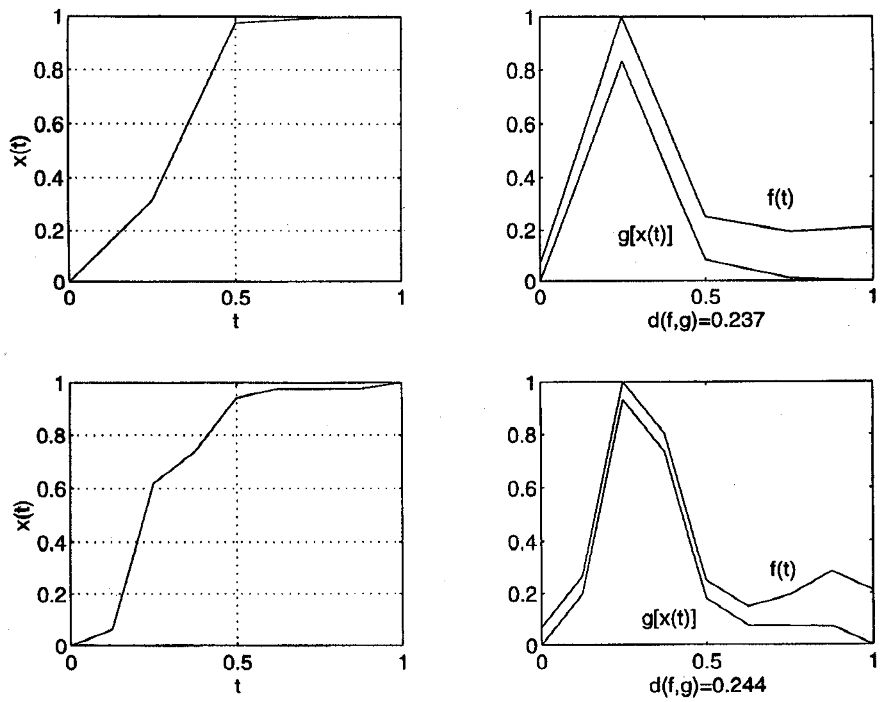


Fig. 6. The output $x(t)$ of the multiresolution algorithm, using $N = 5$ and $N = 9$ samples.

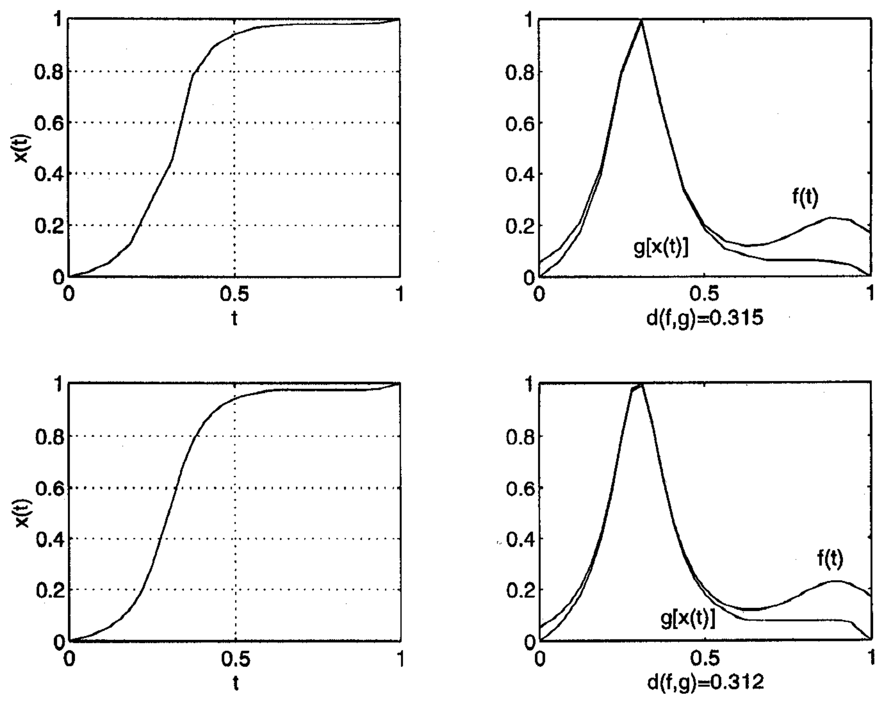


Fig. 7. The output $x(t)$ of the multiresolution algorithm, using $N = 17$ and $N = 33$ samples.

Let f and g be given waveforms. The distance between f and g is defined, in terms of the total variation metric, as

$$d(f,g) = \inf_x \left(0.1 \int |\dot{x} - 1|\, dt + \int \left|\frac{d}{dt}(f(x) - g(t))\right|\right)$$

where the parameters λ_1 and λ_2 have been chosen heuristically. Discretizing t and x into column vectors of length N, the distance $d(f,g)$ is approximated with

$$d(f,g) = \inf_x \Bigg(0.1 \sum_i |(x_i - x_{i-1}) - (t_i - t_{i-1})| + \sum_i |(g(t_i) - f(x_i)) - (g(t_{i-1}) - f(x_{i-1}))|\Bigg)$$

where $f(x_i)$ is obtained using interpolation.

To improve the convergence and speed, a multiresolution approach is adopted. First t and x are discretized into $N = 5$ samples. Then a nonlinear optimization algorithm is run. When the termination tolerances for x and d are met, the number of sampling points is doubled $(N = 2N - 1)$. The algorithm terminates when $N = 33$.

We tested the above algorithm using the functions

$$f(t) = 1/(x-3)^2 + 0.01 + 1/(x-9)^2 + 0.04 - 6$$

and

$$g(t) = \sin(\pi t)$$

displayed in Fig. 5. Fig. 6 shows $x(t)$ for the cases $N = 5$ and $N = 9$, while Fig. 7 shows $x(t)$ for the cases $N = 17$ an $N = 33$.

VII. Observations and Conclusions

Part of the complication in the above analysis arose because we were trying to find a metric that "commutes" between the two types of deformation. The reason is that we wanted a cost function that was invariant with respect to the order of the two different types of noise occurrence, that is, invariant with respect to the incremental deformations: dx_i for domain, and dn_i for amplitude, at time t_i. Note that in the classical amplitude-only deformation case, the invariance was guaranteed by choosing a metric which is a function of the sum of differential amplitudes, or, alternatively, a function of the difference between the transmitted and the received signal. This choice guarantees invariance because, obviously, the sum is invariant with respect to the order of the elements being added. Similarly, in the domain-only deformation case, the invariance was guaranteed by choosing a metric that depends on the overall domain deformation $x(t)$, regardless of the individual domain deformations. Also, by forcing an isometric embedding into the space W, we were able to guarantee that the shortest path of differential deformations dx_i, is actually the line segment connecting the transmitted and received signals in the space W.

When both types of deformations were considered, special care was needed to make sure that, when the different types of deformations intermix, the cost function is not affected. The main result of this paper is that this can be guaranteed if the amplitude metric is a composition-invariant function of $f - g$, and the domain metric is a composition-invariant metric of $x(t) - t$. Examples of such metrics were given.

References

[1] M. A. Akra, "Automated text recognition," Ph.D. dissertation, MIT, Cambridge, MA, 1993.
[2] A. Kolmogorov and S. Fomin. *Elements of the Theory of Functions and Functional Analysis.* New York: Graylock, 1957, p. 36.
[3] J. Milnor. *Morse Theory.* Princeton, NJ: Princeton Univ. Press, 1973, p. 63.
[4] S. T. Rachev. *Probability Metrics and the Stability of Stochastic Models.* New York: Wiley, 1991, p. 8.
[5] I. Shreider. *What is Distance?* Chicago, IL: Univ. Chicago Press, 1974.