Why Inverse F-transform? A Compression-Based Explanation

Vladik Kreinovich
Department of Computer Science
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
vladik@utep.edu

Irina Perfilieva and Vilem Novák
Centre of Excellence IT4Innovations
division of the University of Ostrava
Institute for Research and Applications of Fuzzy Modeling
ul. 30. dubna 22, 701 00 Ostrava 1, Czech Republic
Irina.Perfilieva@osu.cz, Vilem.Novak@osu.cz

1. Data Compression

- In practice, we often need to compress the data:
  - Sometimes, it takes too much space to store all this information.
  - Sometimes, it takes too much computation time to process all this information.
- In these situations, we need to *compress* the data, i.e.:
  - to replace the original values $x(t)$ corresponding to different moments of time t
  - with a few combinations x_i of these values.
- Averaging close values $x(t)$ decreases meas. error, so we take $x_i = \int a_i(t) \cdot x(t)\,dt$, for $t \approx t_i$ for some t_i.
- Optimal $a_i(t)$ should not change with shift, so $a_j(t) = a_i(t - s)$ for some s, hence $a_i(t) = a(t - t_i)$.

2. F-Transform

- We can normalize $a(t)$ by taking $\max a(t) = 1$.
- It is often helpful to assume that $a_i(t) + a_{i+1}(t) = 1$.
- Let t_0 and $h > 0$ be real numbers, let $n > 0$ be an integer, and let $t_i \stackrel{\text{def}}{=} t_0 + i \cdot h$. Let $e(t)$ be s.t.:
  - $e(t) = 0$ for $t \notin (-h, h)$, $e(0) = 1$,
    $e(t) \uparrow$ for $t \in (-h, 0]$, $e(t) \downarrow$ for $t \in [0, h)$, and
  - $e_i(t) + e_{i+1}(t) = 1$ for all $t \in [t_i, t_{i+1}]$, where
    $e_i(t) \stackrel{\text{def}}{=} e(t - t_i)$.
- Then, for each function $x(t)$, its *F-transform* is:

$$x_i = \frac{\int_{t_{i-1}}^{t_{i+1}} e_i(t) \cdot x(t)\, dt}{\int_{t_{i-1}}^{t_{i+1}} e_i(t)\, dt}, \quad i = 1, 2, \ldots, n-1.$$

3. Inverse F-transform

- *Problem:* once we have x_i, we need to reconstruct the original signal.
- *Natural idea:* we have fuzzy rules that if $t \approx t_i$, then $f(t) \approx x_i$.
- We can view $e_i(t) = e(t - t_i)$ as membership functions corresponding to $t \approx t_i$.
- Then, due to $\sum e_i(t) = 1$ for all t, the usual defuzzification leads to

$$\hat{x}(t) = \sum_{i=1}^{n} x_i \cdot e_i(t).$$

- This is known as *inverse F-transform.*
- *Comment:* inverse F-transform is also useful in de-noising, smoothing, etc.

4. At First Glance, Inverse F-Transform Is Counterintuitive

- We want to find $\hat{x}(t)=\sum\limits_{i=0}^{n} c_i\cdot e_i(t)$ which is the closest to $x(t)$: $\displaystyle\int\left(\sum_{i=0}^{n} c_i\cdot e_i(t)-x(t)\right)^2\to\min.$
- Differentiating w.r.t. c_i and equating derivatives to 0, we get: $\displaystyle\sum_{i=1}^{n} c_i\cdot\int e_i(t)\cdot e_j(t)\,dt=\int x(t)\cdot e_j(t)\,dt.$
- For triangular $e(t)$, we get
$$\frac{2}{3}\cdot h\cdot c_j+\frac{1}{6}\cdot h\cdot c_{j-1}+\frac{1}{6}\cdot h\cdot c_{j+1}=x_j.$$
- Clearly, the solution $c_i=x_i$ corresponding to the inverse F-transform does not satisfy these equations.
- Thus, with respect to the above optimization criterion, the inverse F-transform is *not* optimal.

5. A General Definition of an Inverse Transform

- We want a reconstruction of the type $\hat{x}(t) = \sum\limits_{i=0}^{n} c_i \cdot e_i(t)$.
- In inverse F-transform, we take $c_i = x_i$.
- For the optimal least squares approximation, we have a linear system

$$\sum_{i=1}^{n} c_i \cdot \int e_i(t) \cdot e_j(t)\, dt = \int x(t) \cdot e_j(t)\, dt = \left(\int e(t)\, dt\right) \cdot x_i.$$

- The resulting c_i are linear comb. of x_j: $c_i = \sum\limits_j k_{i,j} \cdot x_j$.
- By an *inverse transform*, we mean a matrix K with elements $k_{i,j}$, $0 \le i, j \le n$.
- For each matrix K, by a K*-inverse transform*, we mean a function $\hat{x}_K(t) = \sum\limits_{i=0}^{n} c_i \cdot e_i(t)$, where $c_i = \sum\limits_{j=0}^{n} k_{i,j} \cdot x_j$.

6. A Reasonable Property: Local Consistency

- The main purpose of the F-transform compression is that x_i describe the signal's behavior signal on $[t_i, t_{i+1}]$.
- It is reasonable to require that the inverse transform follows the same idea; e.g.:
  - if the original $x(t)$ signal was constant in a neighborhood of this interval,
  - then $\hat{x}_K(t)$ should also be equal to the same constant for all t from this interval.
- We say that a matrix K is *locally consistent* if
  - for every function $x(t)$ which is equal to a constant c on an interval $[t_i - h, t_{i+1} + h]$,
  - the reconstructed function $\hat{x}_K(t)$ is equal to the same constant c for all $t \in [t_i, t_{i+1}]$.

7. Discussion and Main Result

- One can easily check that the inverse F-transform satisfies this property.
- The F-transform example explains why $\hat{x}_K(t) = c$ only on a subinterval $[t_i, t_{i+1}] \subset [t_i - h, t_{i+1} + h]$:
  - if $x(t) = 1$ for all $t \in [t_i - h, t_{i+1} + h]$ and $x(t) = 0$ for all other t,
  - then inverse F-transform $\hat{x}(t)$ is only equal to 1 for $t \in [t_i, t_{i+1}]$; for all other t, we have $\hat{x}_K(t) < 1$.
- We show that the inverse F-transform is the only locally consistent K-inverse transform.
- **Proposition.** *A matrix K is locally consistent if and only if it coincides with the unit matrix.*
- So, *the above result provides the desired justification of the inverse F-transform.*

8. Proof

- Let $x(t) = 1$ for all $t \in [t_i - h, t_{i+1} + h]$.
- Local consistency implies $\hat{x}_K(t) = 1$ for all $t \in [t_i, t_{i+1}]$.
- For $t \in [t_i, t_{i+1}]$, only $e_i(t)$ and $e_{i+1}(t)$ are non-zero, so $c_i \cdot e_i(t) + c_{i+1} \cdot e_{i+1}(t) = 1$ for all such t.
- For $t = t_i$, we have $e_i(t_i) = 1$, $e_{i+1}(t_i) = 0$, so $c_i = 1$.
- Similarly, for $t = t_{i+1}$, we get $c_{i+1} = 1$.
- Thus, we must have $\sum\limits_{j=0}^{n} k_{i,j} \cdot x_j = 1$ and $\sum\limits_{j=0}^{n} k_{i+1,j} \cdot x_j = 1$.
- By definition of x_i, we get $\sum\limits_{j=0}^{n} k_{i,j} \cdot \int a_j(t) \cdot x(t)\,dt = 1$, i.e., $\int w(t) \cdot x(t)\,dt = 1$, where $w(t) \stackrel{\text{def}}{=} \sum\limits_{j=0}^{n} k_{i,j} \cdot a_j(t)$.
- For $t \notin [t_i - h, t_{i+1} + h]$, $x(t)$ can be arbitrary.

9. Proof (cont-d)

- $\int w(t)\cdot x(t)\,dt = 1$ no matter what the values $x(t)$ for $t \notin [t_i - h, t_{i+1} + h]$.
- Thus, for for $t \notin [t_i - h, t_{i+1} + h]$, we have
$$w(t) = \sum_{j=1}^{n} k_{i,j} \cdot a_j(t) = 0.$$
- For each integer $\ell \le i - 2$, values t from the interval $[t_\ell, t_{\ell+1}]$ are outside the interval $[t_i - h, t_{i+1} + h]$.
- Hence we have $\sum\limits_{j=1}^{n} k_{i,j} \cdot a_j(t) = 0$ for all such t.
- On this interval, only two functions $a_j(t)$ are different from 0: $a_\ell(t)$ and $a_{\ell+1}(t)$, so we get
$$k_{i\ell} \cdot a_\ell(t) + k_{i,\ell+1} \cdot a_{\ell+1}(t) = 0.$$
- In particular, for $t = t_\ell$, we have $a_\ell(t_\ell) \ne 0$ and $a_{\ell+1}(t_\ell) = 0$, so $k_{i,\ell} \cdot a_\ell(t_\ell) = 0$ and thus, $k_{i,\ell} = 0$.

10. Proof (cont-d)

- Similarly, for $t = t_{\ell+1}$, we have $a_\ell(t_{\ell+1}) = 0$ and $a_{\ell+1}(t_{\ell+1}) \neq 0$, so $k_{i,\ell+1} \cdot a_\ell(t_{\ell+1}) = 0$ and thus, $k_{i,\ell+1} = 0$.
- For every integer $\ell \leq i - 2$, we thus have $k_{i,\ell} = 0$ and $k_{i,\ell+1} = 0$. So, we have $k_{i,0} = k_{i,1} = \ldots = k_{i,i-1} = 0$.
- By considering integers $\ell \geq i + 2$, we can similarly get $k_{i,i+2} = \ldots = k_{i,n} = 0$.
- Thus, from $c_i = 1$, it follows that the only possibly non-zero elements $k_{i,j}$ are $k_{i,i}$ and $k_{i,i+1}$.
- Similarly, from $c_{i+1} = 1$, we conclude that the only possibly non-zero elements $k_{i+1,j}$ are $k_{i+1,i}$ and $k_{i+1,i+1}$.
- For $i' = i + 1$, $k_{i',i'-1} = 0$ implies $k_{i+1,i} = 0$.
- Thus, the matrix $k_{i,j}$ is diagonal.
- For a function $x(t) = 1$ for $t \in [t_i - h, t_{i+1} + h]$, the condition $\hat{x}_K(t) = 1$ leads to $k_{i,i} = 1$. Q.E.D.

11. Conclusion

- Empirically, F-transform often provides a good quality compression of signals and images.
- However, it is not optimal w.r.t. standard criteria of compression quality.
- This discrepancy shows that the standard criteria are not fully adequate.
- We propose a new criterion of *local* consistency between the original and the reconstructed signals.
- We show that F-transform is the only scheme that satisfies this criterion.
- Thus, we provide a theoretical justification of the empirical success of F-transform in compression.

12. Acknowledgment

- This work relates to US Dept. of the Navy Grant N62909-12-1-7039 issued by Office of Naval Research Global.
- Additional support was given also:
  - by the European Regional Development Fund in the IT4Innovations Centre of Excellence project CZ.1.05/1.1.00/02.0070,
  - by the US National Science Foundation grants HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and DUE-0926721, and
  - by Grants 1 T36 GM078000-01 and 1R43TR000173-01 from the US National Institutes of Health.