# Overview

- Importance of Features

- Mathematical Notation & Background
- Fourier Transform
- Windowed Fourier Transform
- Wavelets

- Feature Anecdote

- Literature & Homework

# General Remarks

Classifier

Feature vector $(x_1, x_2, \ldots, x_n)$

Feature Extraction

Photograph by MIT OCW.

"The choice of features is more important than the choice of the classifier."

# General Remarks—Application specific

Traffic sign recognition

Color, shape (Hough transform)

Texture Recognition

DFT, WT

Action Recognition

Motion-based features

# General Remarks

Is the choice of features really more important than the choice of the classifier?

We know that a fly uses optical flow features for navigation.

Still, technical systems using optical flow for navigation are (far) behind capabilities of a fly.

# General Remarks—why talk about FT, WFT, WT?

Fourier Transform(FT), Windowed FT (WFT) and Wavelet Transform (WT)

- used in many computer vision applications
- derivation from signal processing
- basic tools for engineers

Other features:
color, motion features (optical flow),
gradient features, SIFT (orientation histograms),
affine invariant features,
steerable filters (overcomplete wavelets),
…

# Background—Notation

| | |
|---|---|
| $\mathbf{Z}, \mathbf{R}, \mathbf{C}$ | integers, real, complex |
| $\mathcal{H}, \mathcal{L}, L^2(\mathbf{R})$ | function spaces |
| $\lVert f \rVert$ | Norm |
| $\lvert a + jb \rvert = \sqrt{a^2 + b^2}$ | Absolute value |
| $\langle f, g \rangle$ | Inner product |
| $\overline{f}(t)$ | Complex conjugate |
| $\hat{f}(\omega)$ | Fourier transform |
| $\cos 2\pi\omega t + j \sin 2\pi\omega t = e^{j2\pi\omega t}$ | Euler formula |

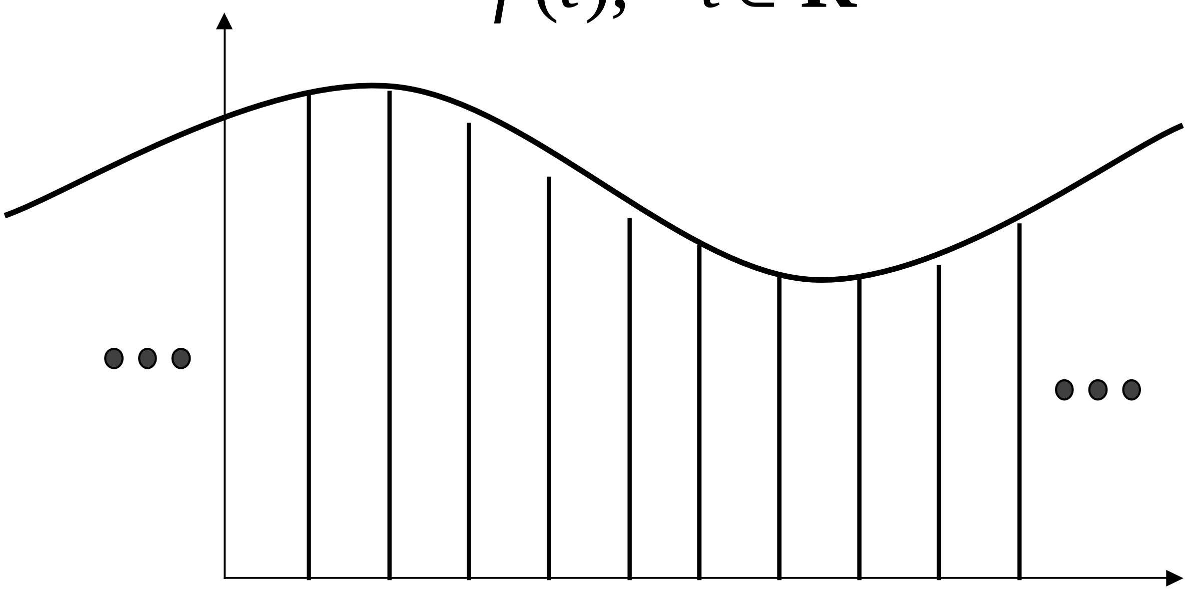
# Background—Vector and Function Spaces

## Vectors and Functions

$$\mathbf{u} = \left[u_1, \ldots, u_N\right]^T$$

$$f(n), \quad n \in \mathbf{Z}$$

$$f(t), \quad t \in \mathbf{R}$$



# Background—Inner Product&Norm

## Inner Product&Norm

$$\langle \mathbf{u}, \mathbf{v} \rangle \equiv \sum_{n=1}^{N} \bar{u}_n v_n \qquad \|\mathbf{u}\|^2 \equiv \langle \mathbf{u}, \mathbf{u} \rangle$$

$$L^2 \text{ norm: } \|\mathbf{u}\|_{L2}^2 = \sum_n |u_n|^2$$

$$\langle f(n), g(n) \rangle \equiv \sum_{-\infty}^{\infty} \bar{f}(n)\ g(n) \qquad \| f(n) \|^2 \equiv \langle f(n), f(n) \rangle$$

$$\langle f(t), g(t) \rangle \equiv \int_{-\infty}^{\infty} \bar{f}(t)\ g(t)\, dt \qquad \| f(t) \|^2 \equiv \langle f(t), f(t) \rangle$$

# Background—Function Spaces

## Inner Product cont.

*Positivity:* $\|f(t)\| > 0$ for all $f \in \mathcal{H},\ f \neq 0$

*Hermiticity:* $\langle f, g \rangle = \overline{\langle g, f \rangle}$

*Linearity:* $\langle f, cg + h \rangle = c\langle f, g \rangle + \langle f, h \rangle$ for $f, g, h \in \mathcal{H}, c \in \mathbf{C}$

## Triangle & Schwarz inequality

$$\|f + g\| \leq \|f\| + \|g\|, \quad |\langle f, g \rangle| \leq \|f\|\,\|g\| \text{ for all } f, g \in \mathcal{H}$$

$L^2(\mathbf{R})$ Function space, finite energy

$$\mathcal{H} \equiv \left\{ f : \mathbf{R} \to \mathbf{C}, \|f\|^2 \equiv \int |f(t)|^2\, dt < \infty \right\}$$

## Basis of a Vector and Function Space

$\{\mathbf{b}_1,\ldots,\mathbf{b}_N\}$ is a basis of $\mathbf{C}^N$ if $\forall\ \mathbf{u} \in\ \mathbf{C}^N$

$$\mathbf{u} = \sum_{n=1}^{N} u_n \mathbf{b}_n, \quad \{u_1,\ldots,u_N\} \text{is unique, } u_n = \langle \mathbf{b}^n, \mathbf{u} \rangle$$

$\{f_1,\ldots,f_N\}$ is a basis of $\mathcal{H}$ if $\forall\ g \in\ \mathcal{H}$

$$g(t) = \sum_{n=1}^{N} c_n f_n(t), \quad \{c_n,\ldots,c_N\} \text{is unique, } c_n = \langle f^n(t), g(t) \rangle$$

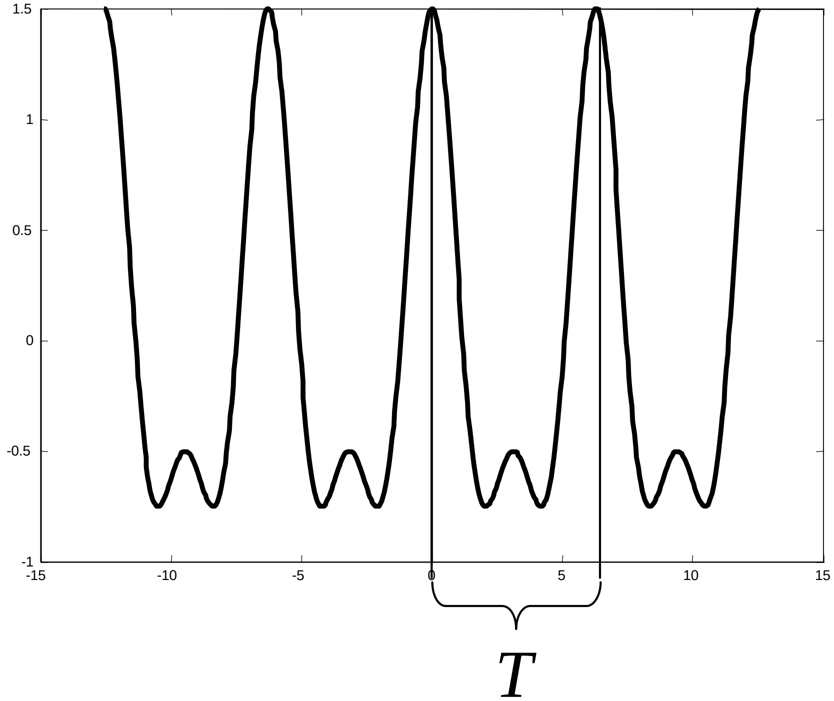
$$g(t) = \int \tilde{g}(\omega) f_\omega(t) d\omega, \quad \tilde{g}(\omega) \text{ is unique, } \tilde{g}(\omega) = \langle f^\omega(t), g(t) \rangle$$

## Orthonormal Basis

$$\langle f_\omega, f_{\omega'} \rangle = \begin{cases} 0 & \omega \neq \omega' \\ 1 & \omega = \omega' \end{cases} \quad f^\omega = f_\omega \quad \tilde{g}(\omega) = \langle f_\omega, g \rangle$$

# Fourier Series—Continuous Signal

## Continuous, periodic signal




$f(t)$ periodic with period $T$

$f(t) \in L^2\left(\left[-T/2, T/2\right]\right)$

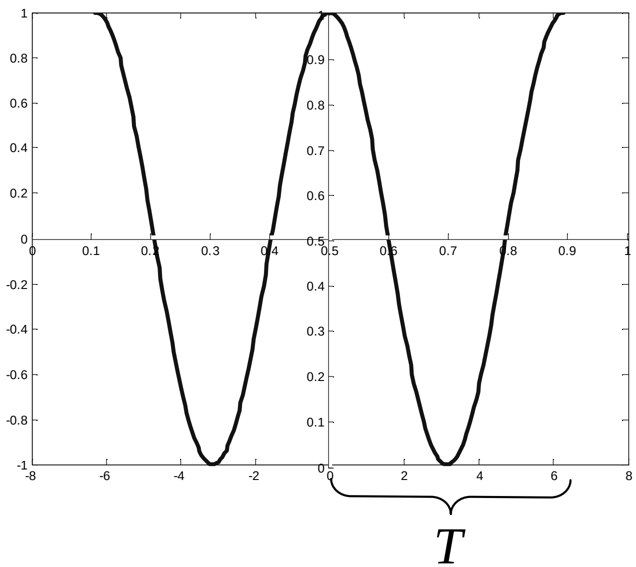
$$c_k = \int_{-T/2}^{T/2} f(t) e^{\frac{-j2\pi kt}{T}} dt$$

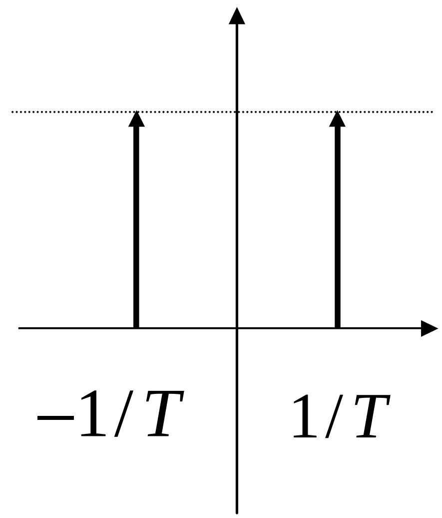
$$s_k = e^{\frac{j2\pi kt}{T}}, \langle s_k, s_l \rangle_T = T\delta_k^l$$

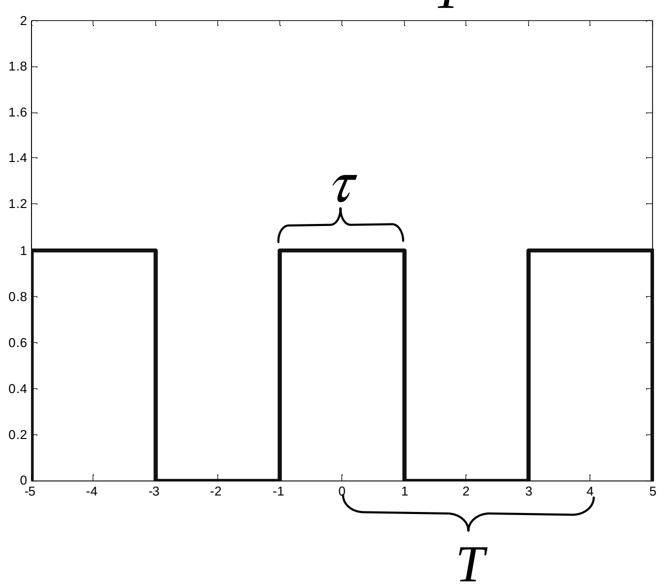
$$c_k = \langle s_k, f \rangle_T, \| f(t) \|_T^2 = \frac{1}{T} \sum_{k=-\infty}^{\infty} |c_k|^2$$

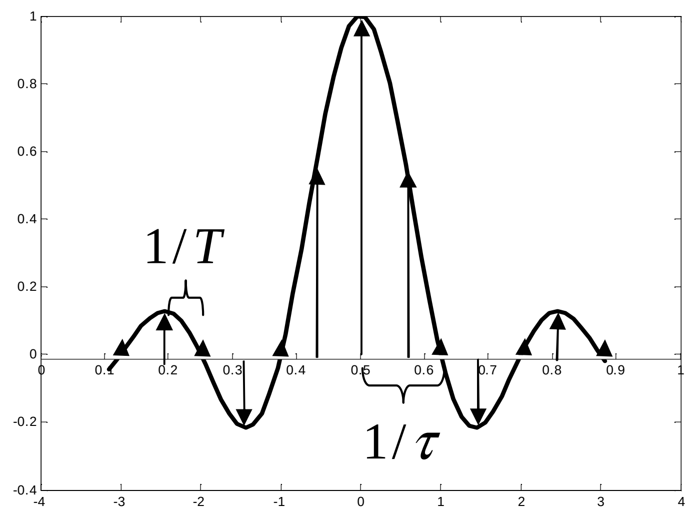
$$f(t) = \frac{1}{T} \sum_{k=-\infty}^{\infty} c_k \, e^{\frac{j2\pi kt}{T}}$$

# Fourier Series—Examples



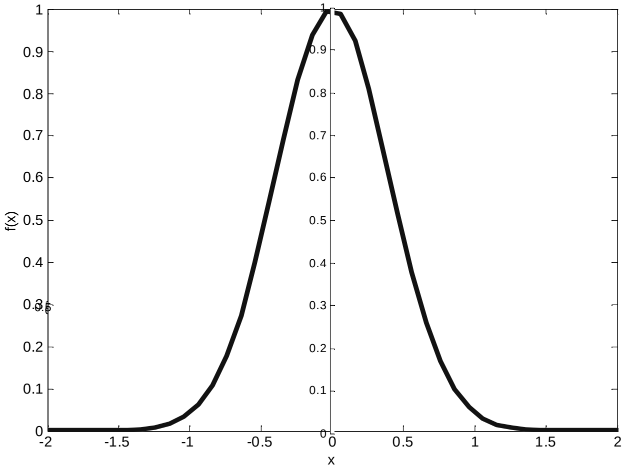



# Fourier Transform

Continuous signals, $f(t) \in L^2(\mathbf{R})$




$$T \to \infty$$

$$L^2\left(\left[-T/2, T/2\right]\right) \to L^2(\mathbf{R})$$

$$\omega_k = k/T$$

$$\Delta\omega = \omega_{k+1} - \omega_k = \frac{1}{T}$$

$$c_k = \int_{-T/2}^{T/2} f(t) e^{-j2\pi\omega_k t} dt \to \hat{f}(\omega) = \int_{-\infty}^{\infty} f(t) e^{-j2\pi\omega t} dt$$

$$f(t) = \frac{1}{T} \sum_{k=-\infty}^{\infty} c_k \, e^{\frac{j2\pi kt}{T}} = \sum_{k=-\infty}^{\infty} c_k \, e^{j2\pi\omega_k} \Delta\omega_k$$

$$T \to \infty \quad f(t) = \int_{-\infty}^{\infty} \hat{f}(\omega) e^{j2\pi\omega t} d\omega$$

Another look at the inverse Fourier transform

$$f(t) = \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} f(t^{'}) e^{-j2\pi\omega t^{'}} dt^{'} \right] e^{j2\pi\omega t} d\omega$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(t^{'}) e^{-j2\pi\omega(t^{'}-t)} d\omega dt^{'} = \int_{-\infty}^{\infty} f(t^{'}) \delta(t - t^{'}) dt^{'}$$

# Fourier Transform—Properties

| | | |
|---|---|---|
| Linearity: | $A f(t) + B g(t)$ | $A \hat{f}(\omega) + B \hat{g}(\omega)$ |
| Shift: | $f(t - t_0)$ | $e^{-j 2\pi \omega t_0} \hat{f}(\omega)$ |
| Convolution: | $\int_{-\infty}^{\infty} f(t') g(t - t') dt'$ | $\hat{f}(\omega)\, \hat{g}(\omega)$ |
| Derivative: | $\frac{d f(t)}{dt}$ | $j 2\pi \omega \hat{f}(\omega)$ |
| Scaling: | $f(At)$ | $\frac{1}{\|A\|} \hat{f}\left(\frac{\omega}{A}\right)$ |
| Correlation : | $\int_{-\infty}^{\infty} \overline{f}(t') g(t + t') dt'$ | $\overline{\hat{f}}(\omega)\, \hat{g}(\omega)$ |
| Autocorrelation: | $\int_{-\infty}^{\infty} \overline{f}(t') f(t + t') dt'$ | $\left\| \hat{f}(\omega) \right\|^2$ |

# Fourier Transform—Plancherel&Parseval

## Plancherel's Theorem

$$\int_{-\infty}^{\infty} \left| f \right|^2 dt = \int_{-\infty}^{\infty} \left| \hat{f} \right|^2 d\omega$$

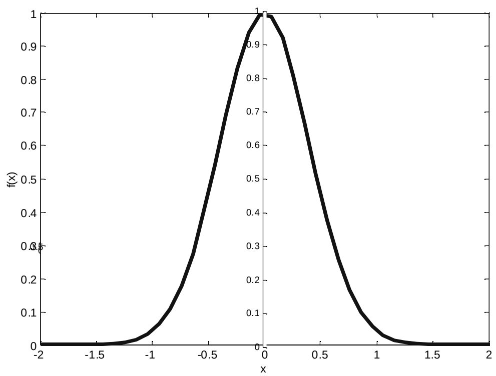
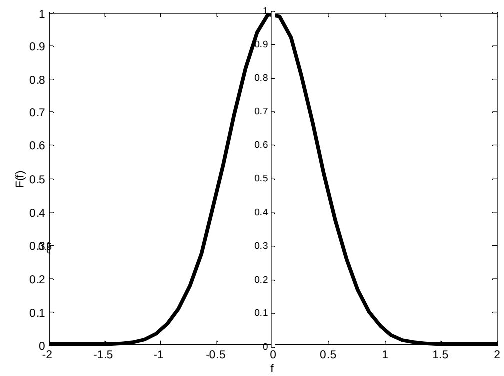
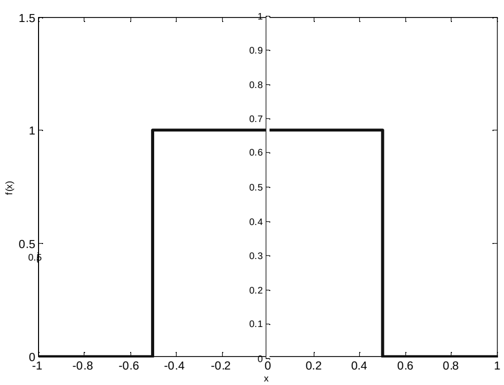
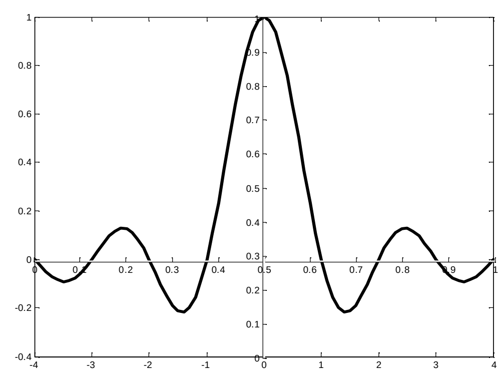
$$\left\| f \right\|^2 = \left\| \hat{f} \right\|^2$$

## Parseval Identity

$$\int_{-\infty}^{\infty} \overline{f}\, g \, dt = \int_{-\infty}^{\infty} \overline{\hat{f}}\, \hat{g} \, d\omega$$

$$\left\langle f, g \right\rangle = \left\langle \hat{f}, \hat{g} \right\rangle$$

# Fourier Transform—Examples

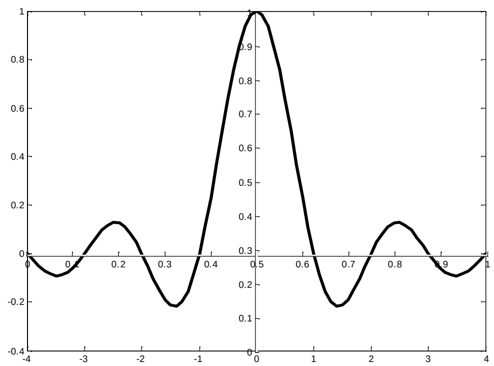
| Time | Frequency |
|---|---|
| $e^{-t^2/(2\sigma^2)}$ | $\sigma\sqrt{2\pi}\,e^{-2\pi^2\sigma^2\omega^2}$ |
|  |  |
| $rect(t/T)$ | $T\dfrac{\sin(\pi\omega T)}{\pi\omega T} = T\ \mathrm{sinc}(\omega T)$ |
|  |  |

# Fourier Transform—Examples

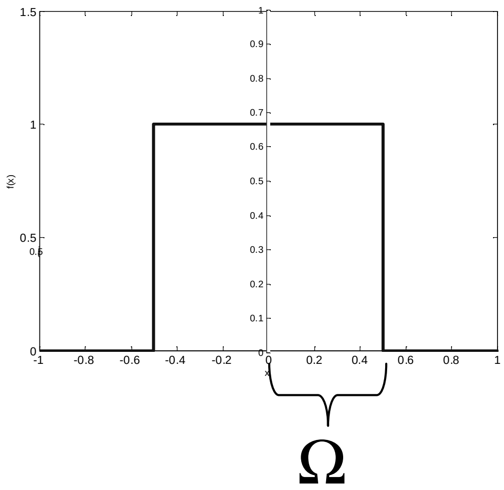
Time

Frequency

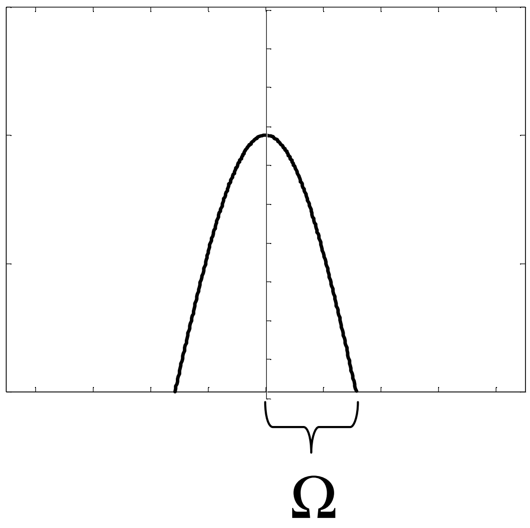
$$2\Omega\frac{\sin(2\pi\Omega t)}{2\pi\Omega t} = 2\Omega\,\mathrm{sinc}(2\Omega t)$$

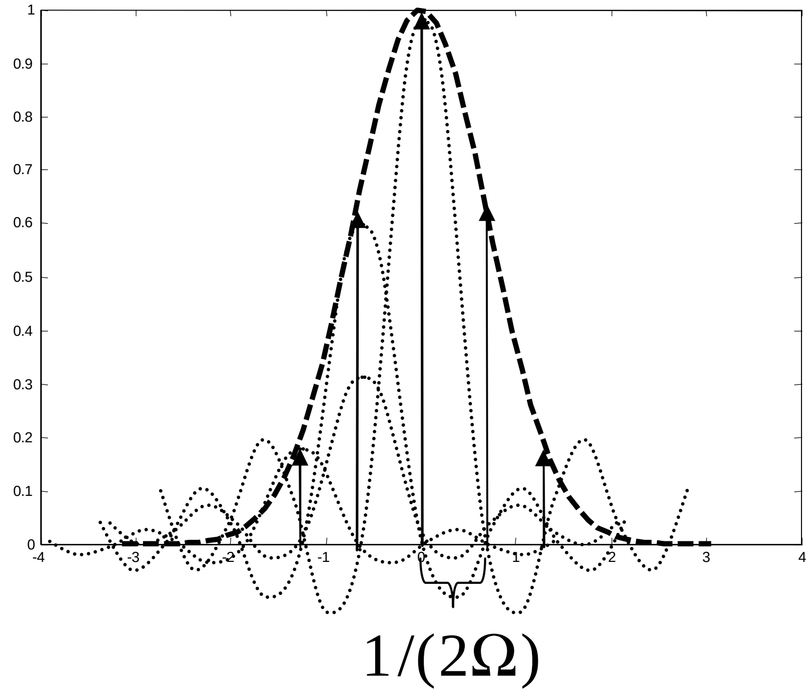


$$rect(\omega/(2\Omega))$$



# Fourier Transform—Sampling Theorem




$$\hat{f}(\omega) = 0, \text{ for } |\omega| > \Omega$$

$$f(t) = \sum_{n=-\infty}^{\infty} \operatorname{sinc}(2\Omega t - n)\, f\left(\frac{n}{2\Omega}\right)$$

# Fourier Transform—Sampling Theorem

$$\hat{f}(\omega) = 0, \text{ for } |\omega| > \Omega, \; f(t) = \sum_{n=-\infty}^{\infty} \operatorname{sinc}(2\Omega t - n)\, f\left(\frac{n}{2\Omega}\right)$$

$$t_n = \frac{n}{2\Omega}, \; f(t) = \sum_{n=-\infty}^{\infty} \operatorname{sinc}(2\Omega(t - t_n))\, f(t_n)$$

expand $\hat{f}(\omega)$ into a Fourier series:

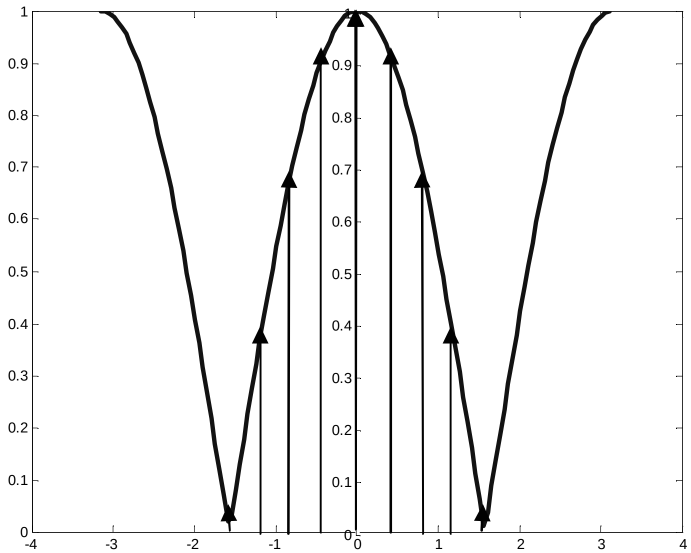
$$\hat{f}(\omega) = \frac{1}{2\Omega} \sum c_n e^{-j2\pi\omega t_n}, \; c_n = \int_{-\Omega}^{\Omega} \hat{f}(\omega) e^{j2\pi\omega t_n} d\omega = \underbrace{\int_{-\infty}^{\infty} \hat{f}(\omega) e^{j2\pi\omega t_n} d\omega}_{\text{Inverse FT}} = f(t_n)$$

$$f(t) = \int_{-\infty}^{\infty} \hat{f}(\omega) e^{j2\pi\omega t} d\omega = \int_{-\Omega}^{\Omega} \hat{f}(\omega) e^{j2\pi\omega t} d\omega$$

$$= \int_{-\Omega}^{\Omega} \frac{1}{2\Omega} \sum f(t_n) e^{j2\pi\omega(t - t_n)} d\omega = \sum f(t_n) \underbrace{\frac{1}{2\Omega} \int_{-\Omega}^{\Omega} e^{j2\pi\omega(t - t_n)} d\omega}_{\text{Inverse FT of rect function}}$$

$$= \sum_{n=-\infty}^{\infty} \operatorname{sinc}(2\Omega(t - t_n))\, f(t_n) \quad \text{basis of bandlimited functions}$$

## Discrete Fourier Transform (DFT)




Discrete, periodic signal

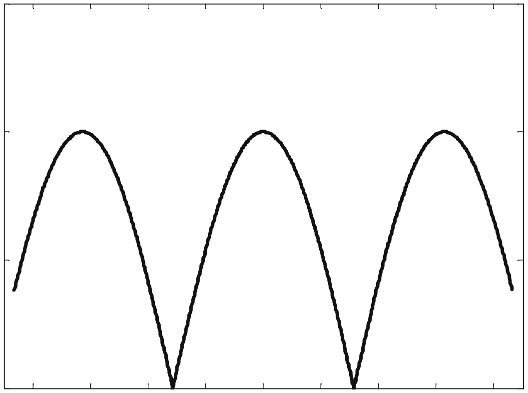
$$c_k = \sum_{n=0}^{N-1} f(n)\, e^{\frac{-j2\pi kn}{N}}$$

$$f(n) = \frac{1}{N} \sum_{k=0}^{N-1} c_k\, e^{\frac{j2\pi kn}{N}}$$

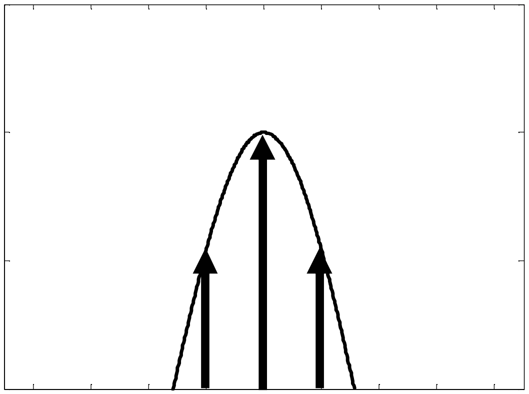
# Fourier Transform—Summary

Time | Frequency

Continuous, periodic

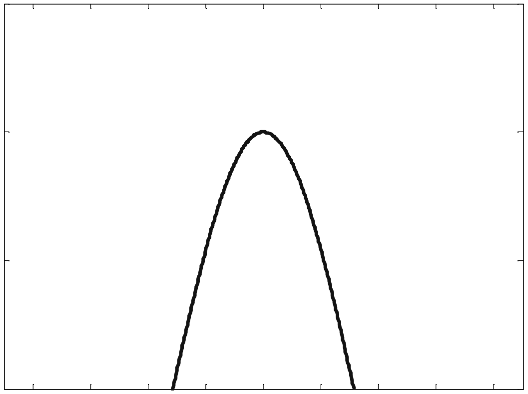


FS

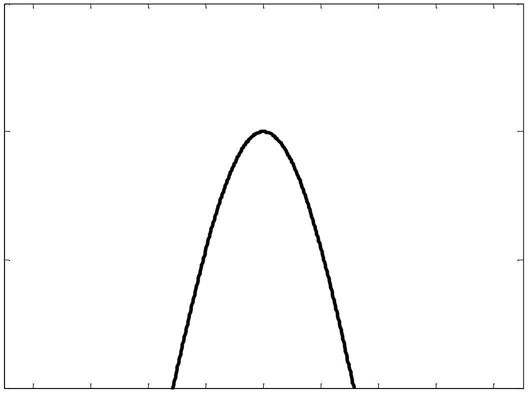


Discrete

Continuous

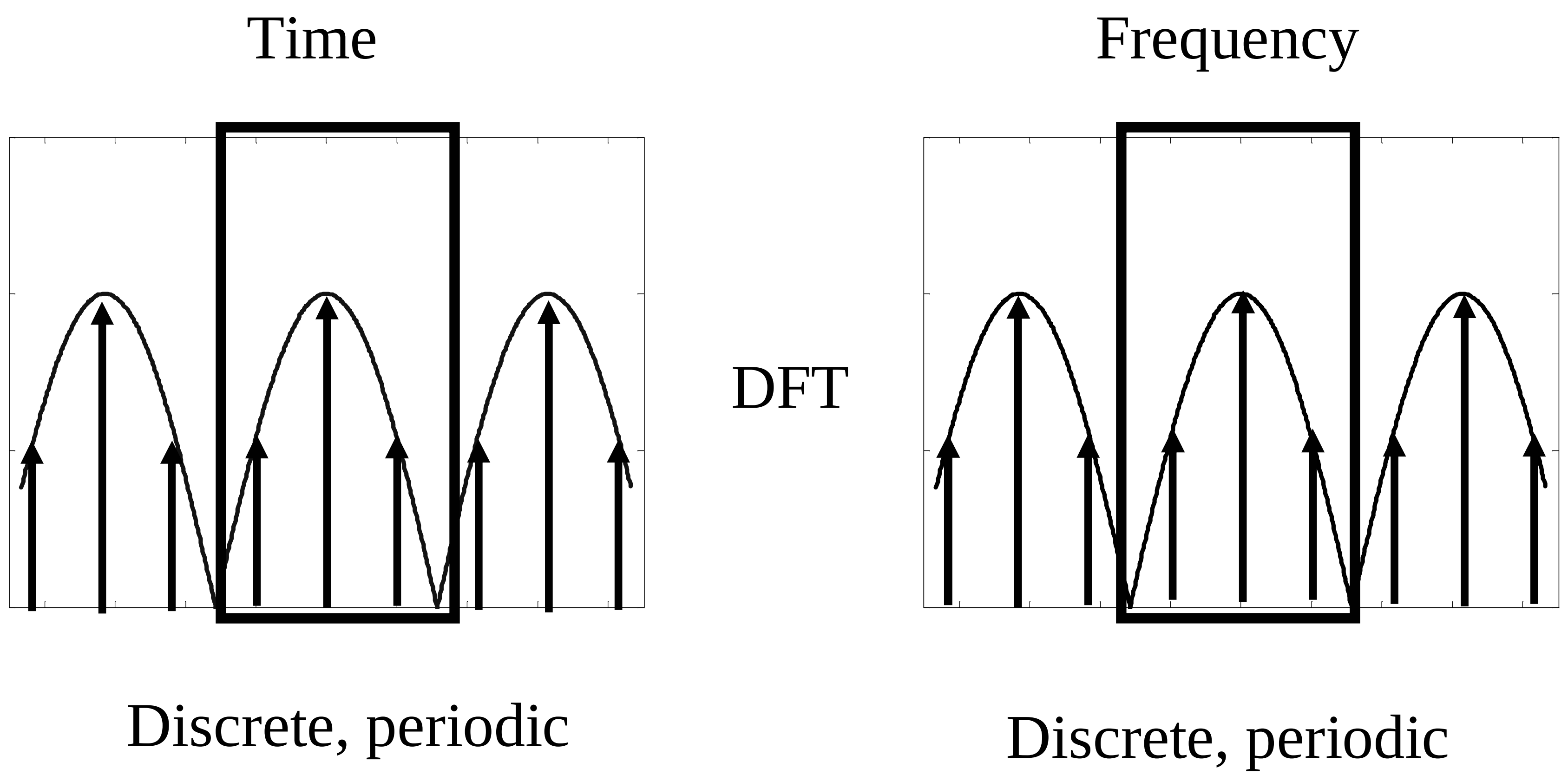


FT



Continuous

# Fourier Transform—Summary



# Discrete Fourier Transform—Fast Fourier Transform

## Decimation in Space

$$c_k = \sum_{n=0}^{N-1} f(n) e^{\frac{-j2\pi kn}{N}} \qquad O(N^2)$$

$$= \sum_{n=0}^{N/2-1} f(2n) e^{\frac{-j2\pi k 2n}{N}} + \sum_{n=0}^{N/2-1} f(2n+1) e^{\frac{-j2\pi k(2n+1)}{N}}$$

$$= \underbrace{\sum_{n=0}^{N/2-1} f(2n) e^{\frac{-j2\pi kn}{N/2}}}_{c_k^e} + \underbrace{e^{\frac{-j2\pi k}{N}}}_{W_k} \underbrace{\sum_{n=0}^{N/2-1} f(2n+1) e^{\frac{-j2\pi kn}{N/2}}}_{c_k^o}$$

$$e^{\frac{-j2\pi k}{N}} = -e^{\frac{-j2\pi(k+N/2)}{N}} \Leftrightarrow W_k = -W_{k+N/2}$$

$$\sum_{n=0}^{N/2-1} f(2n) e^{\frac{-j2\pi kn}{N/2}} = \sum_{n=0}^{N/2-1} f(2n) e^{\frac{-j2\pi(k+N/2)n}{N/2}} \Leftrightarrow c_k^e = c_{k+N/2}^e$$

$$\sum_{n=0}^{N/2-1} f(2n+1) e^{\frac{-j2\pi kn}{N/2}} = \sum_{n=0}^{N/2-1} f(2n+1) e^{\frac{-j2\pi(k+N/2)n}{N/2}} \Leftrightarrow c_k^o = c_{k+N/2}^o$$

# Discrete Fourier Transform— Fast Fourier Transform

$$W_k = -W_{k+N/2}$$

$$c_k^e = c_{k+N/2}^e$$

$$c_k^o = c_{k+N/2}^o$$

$$c_k = \sum_{n=0}^{N-1} f(n) e^{\frac{-j2\pi kn}{N}}$$

$$c_k^e = \sum_{n=0}^{N'-1} f(2n) e^{\frac{-j2\pi kn}{N'}}$$

$$c_k^o = \sum_{n=0}^{N'-1} f(2n+1) e^{\frac{-j2\pi kn}{N'}}$$

$$N' = N/2 \quad O(N^2/2)$$

$$\left.\begin{aligned} c_k &= c_k^e + W_k c_k^o \\ c_{k+N/2} &= c_k^e - W_k c_k^o \end{aligned}\right\} 0 \le k < N/2$$

Recursion:

$$\hat{c}_0 = \hat{c}_0^e + \hat{c}_0^o$$

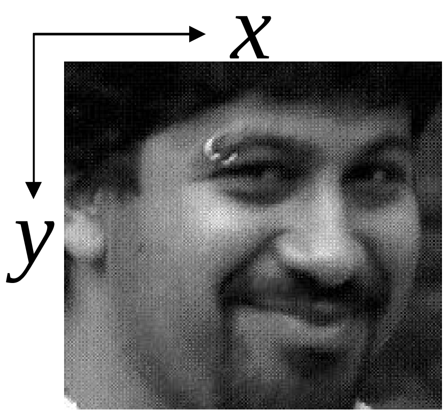
$$\hat{c}_1 = \hat{c}_0^e - \hat{c}_0^o$$

Complexity:

$M/2\ \mathrm{ld}(M)$ Multiplications

$M\ \mathrm{ld}(M)$ Summations

# Discrete Fourier Transform—Image Analysis




Courtesy of Professors Tomaso Poggio and Sayan Mukherjee. Used with permission.

$$c(k_x, k_y) = \frac{1}{MN} \sum_{y=0}^{M-1} \sum_{x=0}^{N-1} f(x,y) e^{\frac{-j2\pi k_x x}{N}} e^{\frac{-j2\pi k_y y}{M}}$$

$$= \frac{1}{MN} \sum_{y=0}^{M-1} \left[ \sum_{x=0}^{N-1} f(x,y) e^{\frac{-j2\pi k_x x}{N}} \right] e^{\frac{-j2\pi k_y y}{M}}$$

*M* 1-D DFT's rows

$$c(k_x, y) = \sum_{x=0}^{N-1} f(x,y) e^{\frac{-j2\pi k_x x}{N}}$$

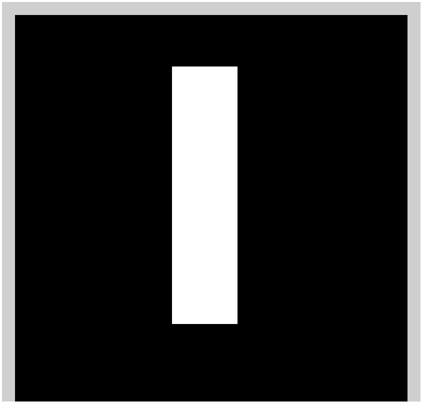
*N* 1-D DFT's columns

$$c(k_x, k_y) = \sum_{y=0}^{M-1} c(k_x, y) e^{\frac{-j2\pi k_y y}{M}}$$

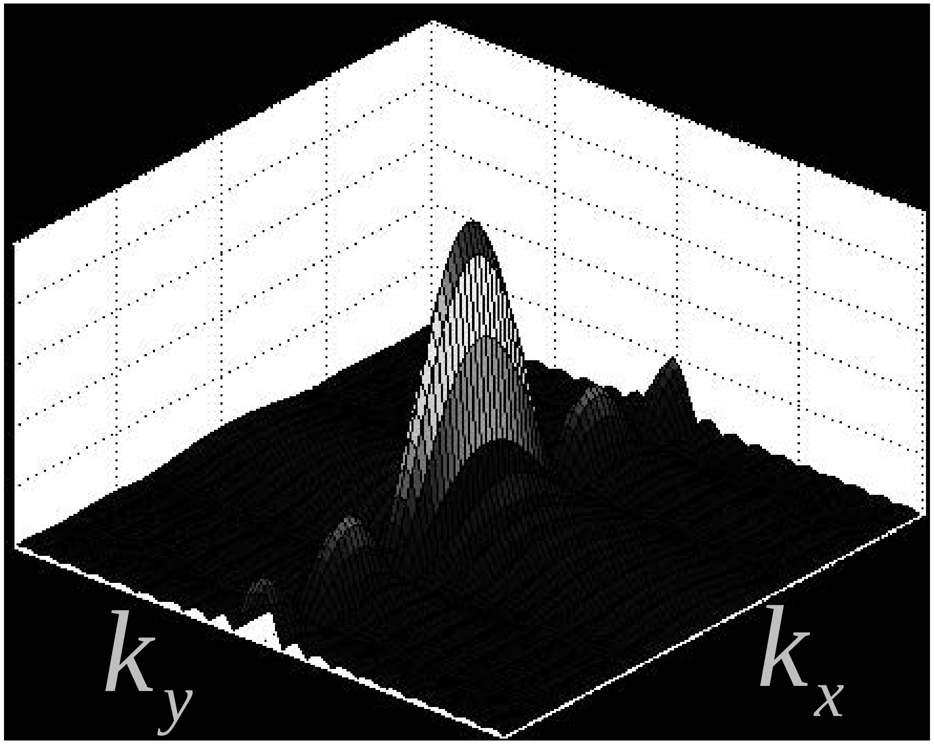
# Discrete Fourier Transform—Image Analysis

## Example

$y$



$x$

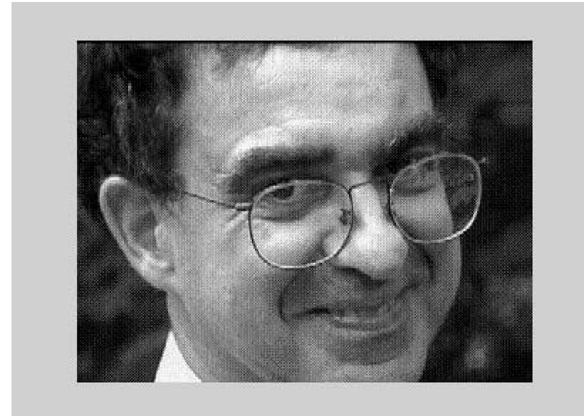


# Discrete Fourier Transform—Image Analysis

Image

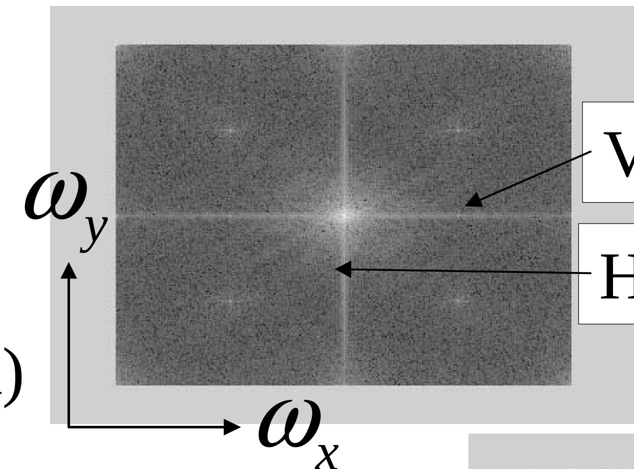
(A)


(B)

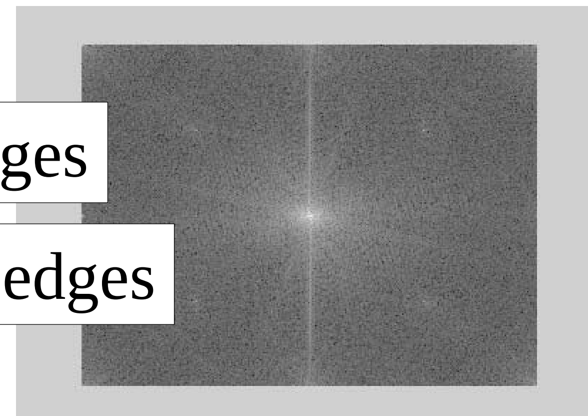


Spectrum

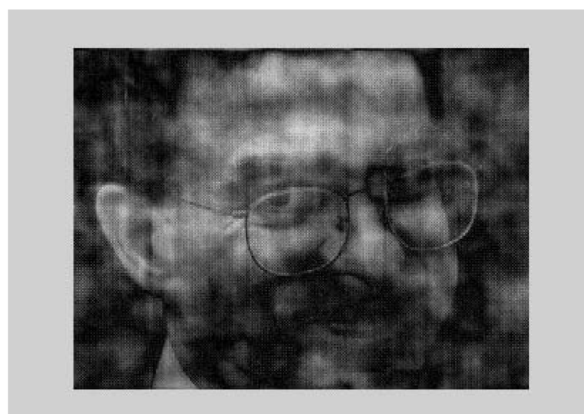
(A)




(B)



Amplitude of (A) &
Phase of (B)

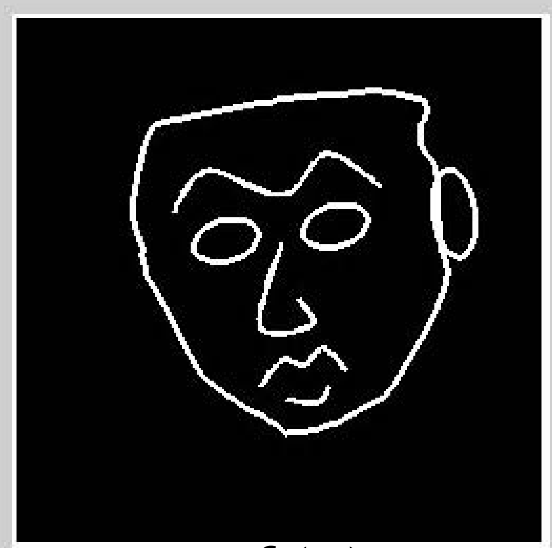


# Discrete Fourier Transform—Image Analysis

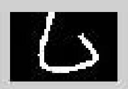
## Template Matching

$$corr(x) = \int_{-\infty}^{\infty} f(x')g(x'-x)dx' \qquad g'(x) = g(-x)$$

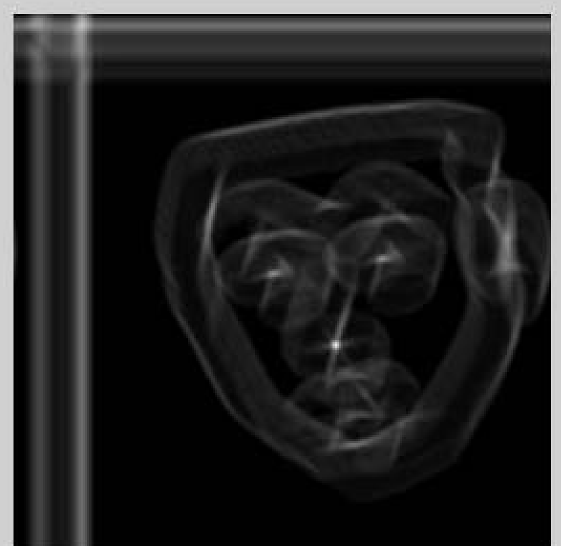
$$corr(x) = f * g' = \int_{-\infty}^{\infty} f(x')g'(x-x')dx' \qquad \hat{f}(\omega)\hat{g}'(\omega)$$



$f(x)$

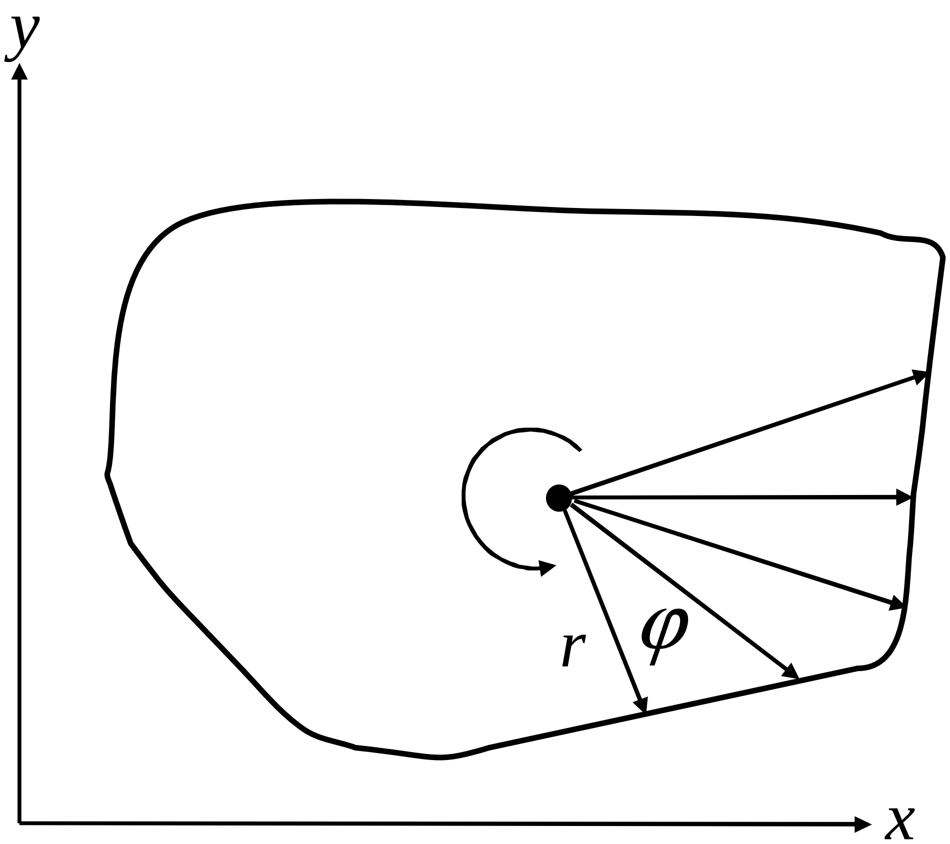


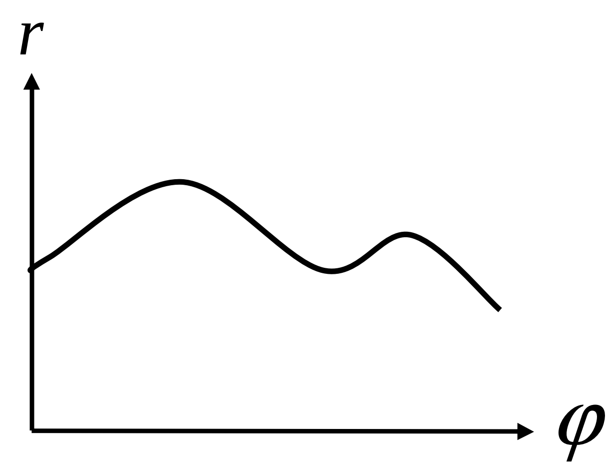
$g(x)$



$corr(x)$

## Shape description with Fourier descriptors

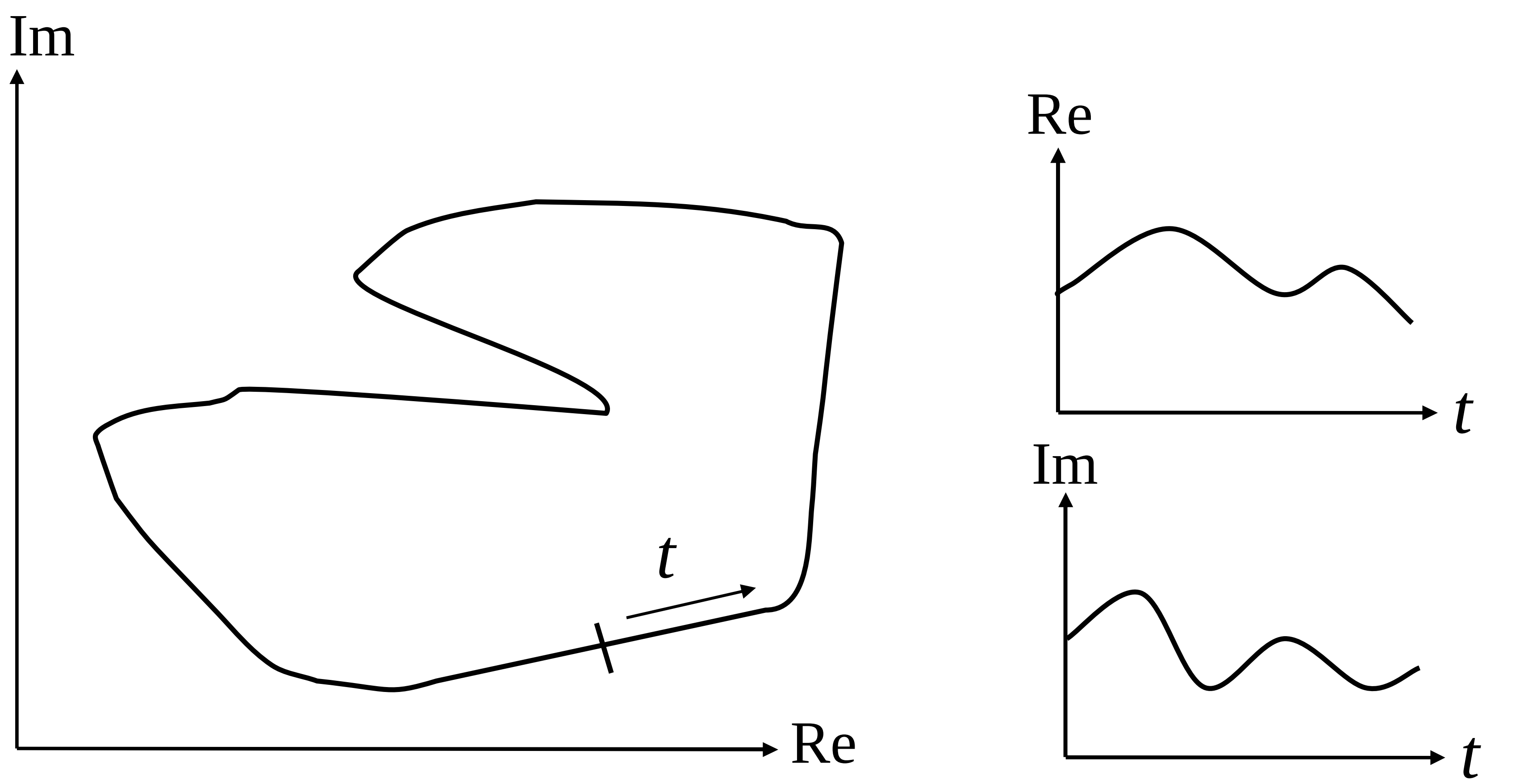



Invariance:
- Scale
- Rotation
- Translation

# Discrete Fourier Transform—Image Analysis

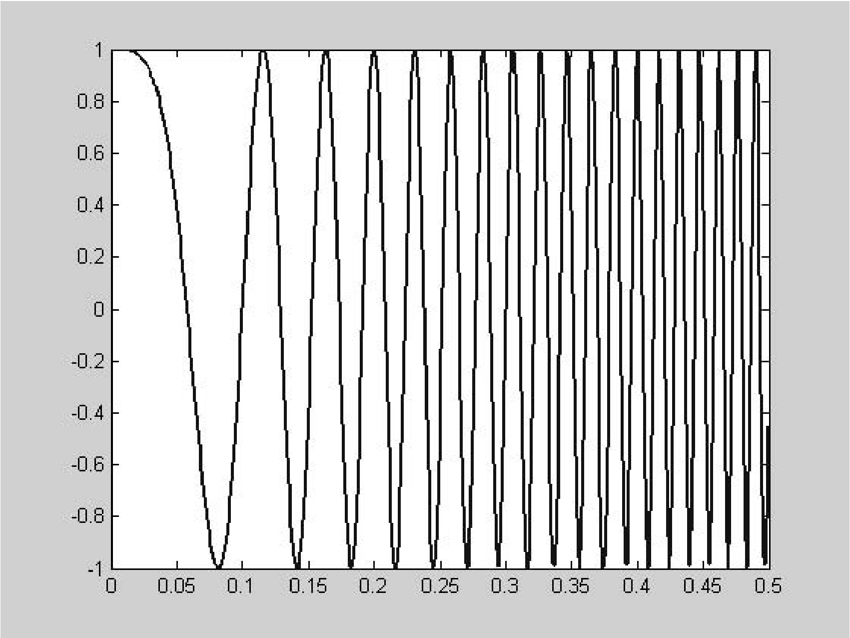
## Shape description with Fourier descriptors

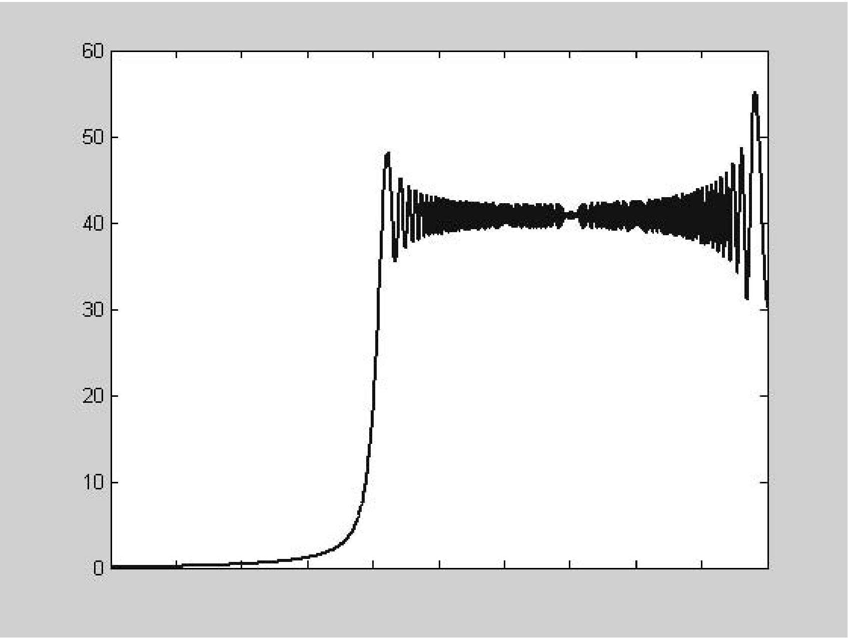


# Windowed Fourier Transform—Motivation

## Fourier transform of a chirp signal

$$\cos(\pi t^2), \omega_{\text{inst}} = t$$



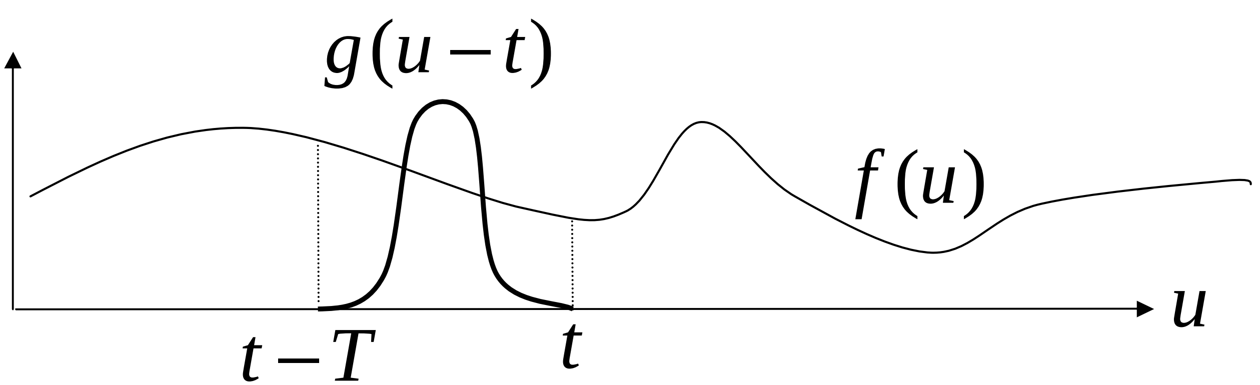
# Windowed Fourier Transform

## Windowed Fourier Transform (WFT)

$$\tilde{f}(\omega,t)=\int_{-\infty}^{\infty} f(u)\overline{g}(u-t)e^{-j2\pi\omega u}du \qquad \text{supp } g \subset [-T,0]$$

$$f_t(u)=f(u)\overline{g}(u-t) \qquad \text{supp } f_t \subset [t-T,t]$$

Fourier transform: $\tilde{f}(\omega,t)=\int_{-\infty}^{\infty} f_t(u)e^{-j2\pi\omega u}du$



# Windowed Fourier Transform—Time Frequency Symmetry

$$\tilde{f}(\omega,t) = \int_{-\infty}^{\infty} f(u)\overline{g}(u-t)e^{-j2\pi\omega u}du$$

substitute $g(u-t)e^{j2\pi\omega u}$ by $g_{\omega,t}(u)$

Parseval's Identity

$$\tilde{f}(\omega,t) = \int_{-\infty}^{\infty} \overline{g}_{\omega,t}(u) f(u)du = \left\langle g_{\omega,t}, f \right\rangle = \left\langle \hat{g}_{\omega,t}, \hat{f} \right\rangle$$

$$\hat{g}_{\omega,t}(\nu) = \int_{-\infty}^{\infty} g(u-t)e^{j2\pi\omega u}e^{-j2\pi\nu u}du = \int_{-\infty}^{\infty} g(u-t)e^{-j2\pi u(\nu-\omega)}du,$$

substitute $u'$ by $u-t$ :

$$\int_{-\infty}^{\infty} g(u')e^{-j2\pi(u'+t)(\nu-\omega)}du' = e^{-j2\pi t(\nu-\omega)}\hat{g}(\nu-\omega)$$

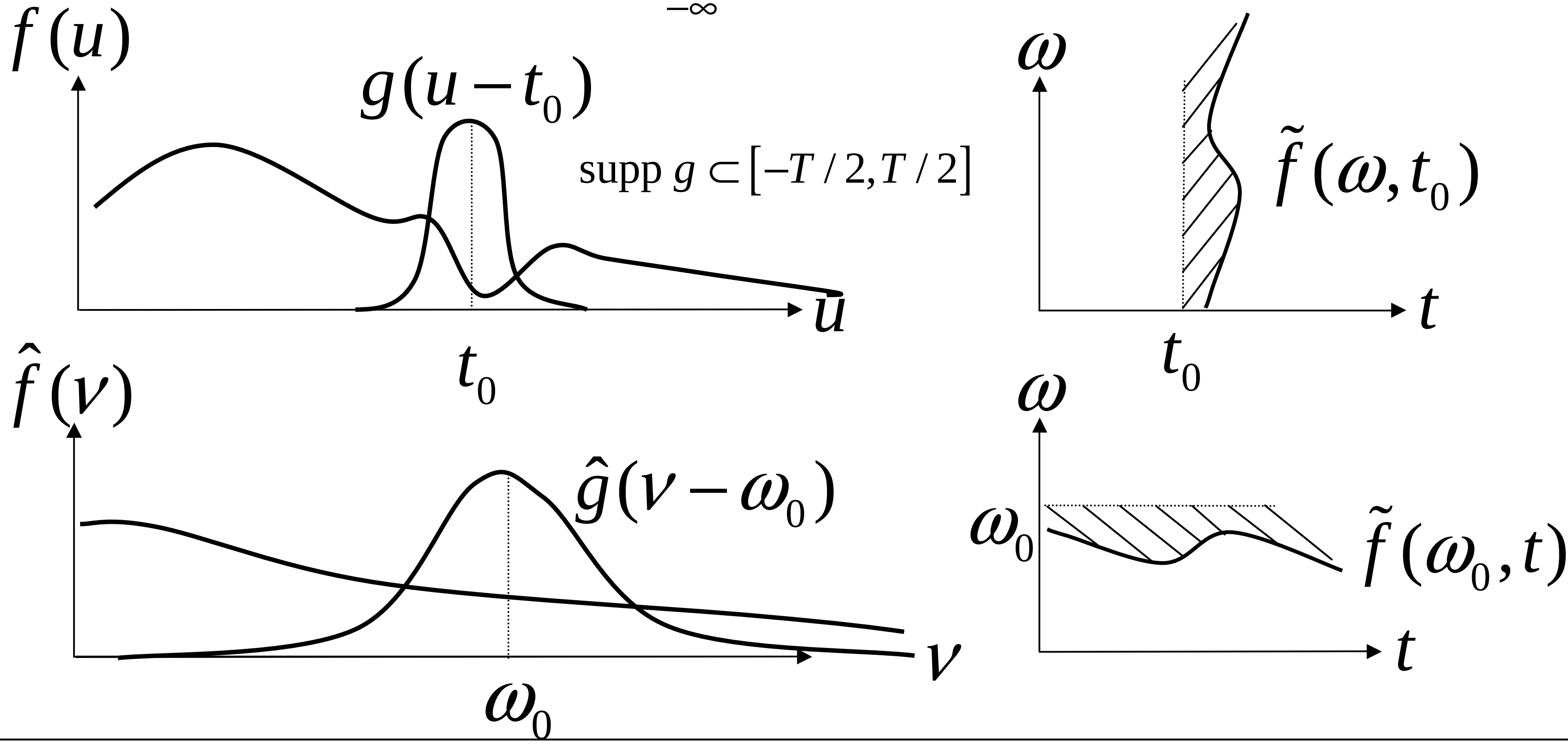
$$\tilde{f}(\omega,t) = e^{-j2\pi\omega t}\int_{-\infty}^{\infty} \overline{\hat{g}}(\nu-\omega)\hat{f}(\nu)e^{j2\pi\nu t}d\nu$$

# Windowed Fourier Transform—Time Frequency Localization

## Time Frequency Symmetry

$$\tilde{f}(\omega,t) = \int_{-\infty}^{\infty} f(u)\overline{g}(u-t)e^{-j2\pi\omega u}du$$

$$\tilde{f}(\omega,t) = e^{-j2\pi\omega t}\int_{-\infty}^{\infty} \overline{\hat{g}}(\nu-\omega)\hat{f}(\nu)e^{j2\pi\nu t}d\nu$$



# Windowed Fourier Transform—Time Frequency Localization

## Time Frequency Localization

$$\|g(t)\|^2 = 1 \qquad \|\hat{g}(\omega)\|^2 = 1$$

$$t_m = \int_{-\infty}^{\infty} t\,|g(t)|^2\,dt \qquad \omega_m = \int_{-\infty}^{\infty} \omega\,|\hat{g}(\omega)|^2\,d\omega$$

$$\sigma_t^{\ 2} = \int_{-\infty}^{\infty} (t - t_m)^2\,|g(t)|^2\,dt \qquad \sigma_\omega^{\ 2} = \int_{-\infty}^{\infty} (\omega - \omega_m)^2\,|\hat{g}(\omega)|^2\,d\omega$$

## Heisenberg's uncertainty principle

$$4\pi\sigma_\omega\sigma_t \geq 1$$

$$g(t) = (2a)^{1/4} e^{-\pi a t^2}$$

$$\hat{g}(\omega) = (2/a)^{1/4} e^{-\pi\omega^2/a}$$

$$t_m = \omega_m = 0$$

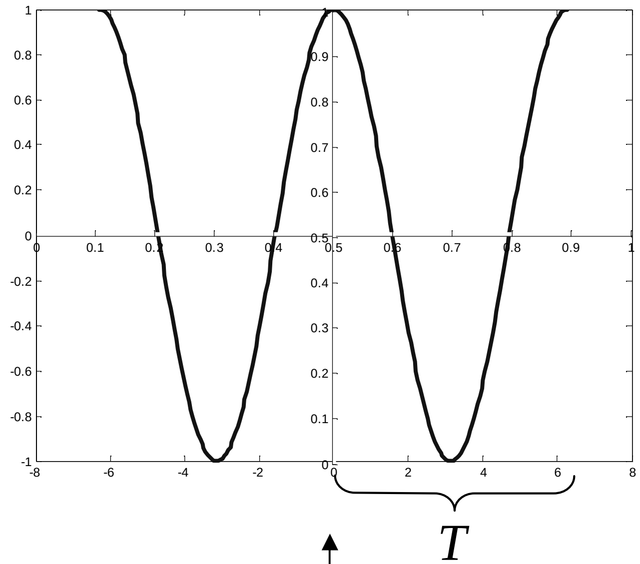
$$\sigma_t = \sqrt{\frac{1}{4\pi a}}$$

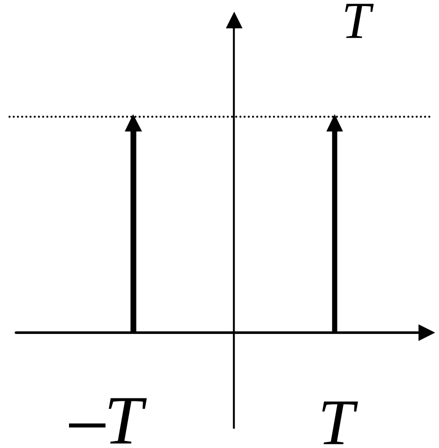
$$\sigma_\omega = \sqrt{\frac{a}{4\pi}}$$

# Windowed Fourier Transform—Time Frequency Localization

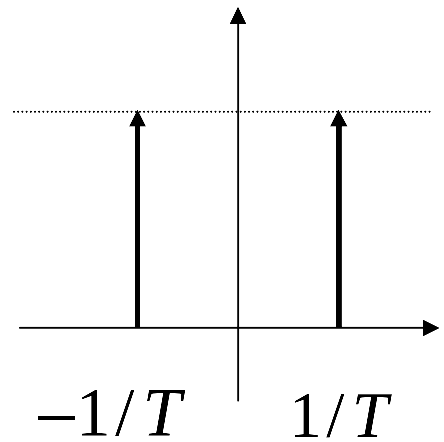
## Uncertainty Principle

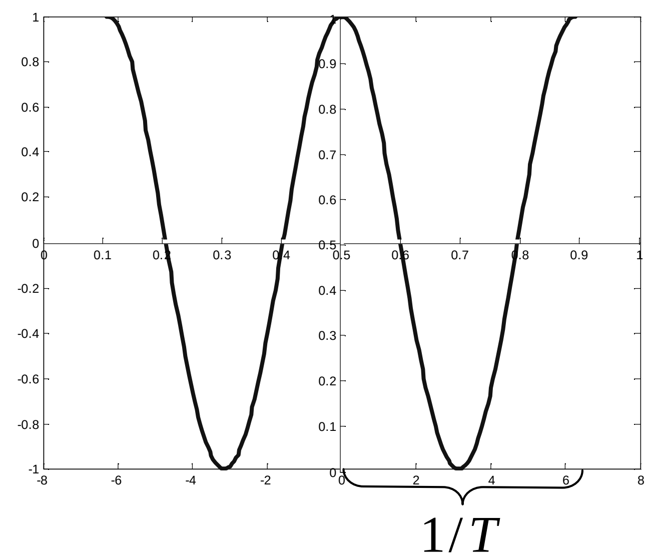
### Time



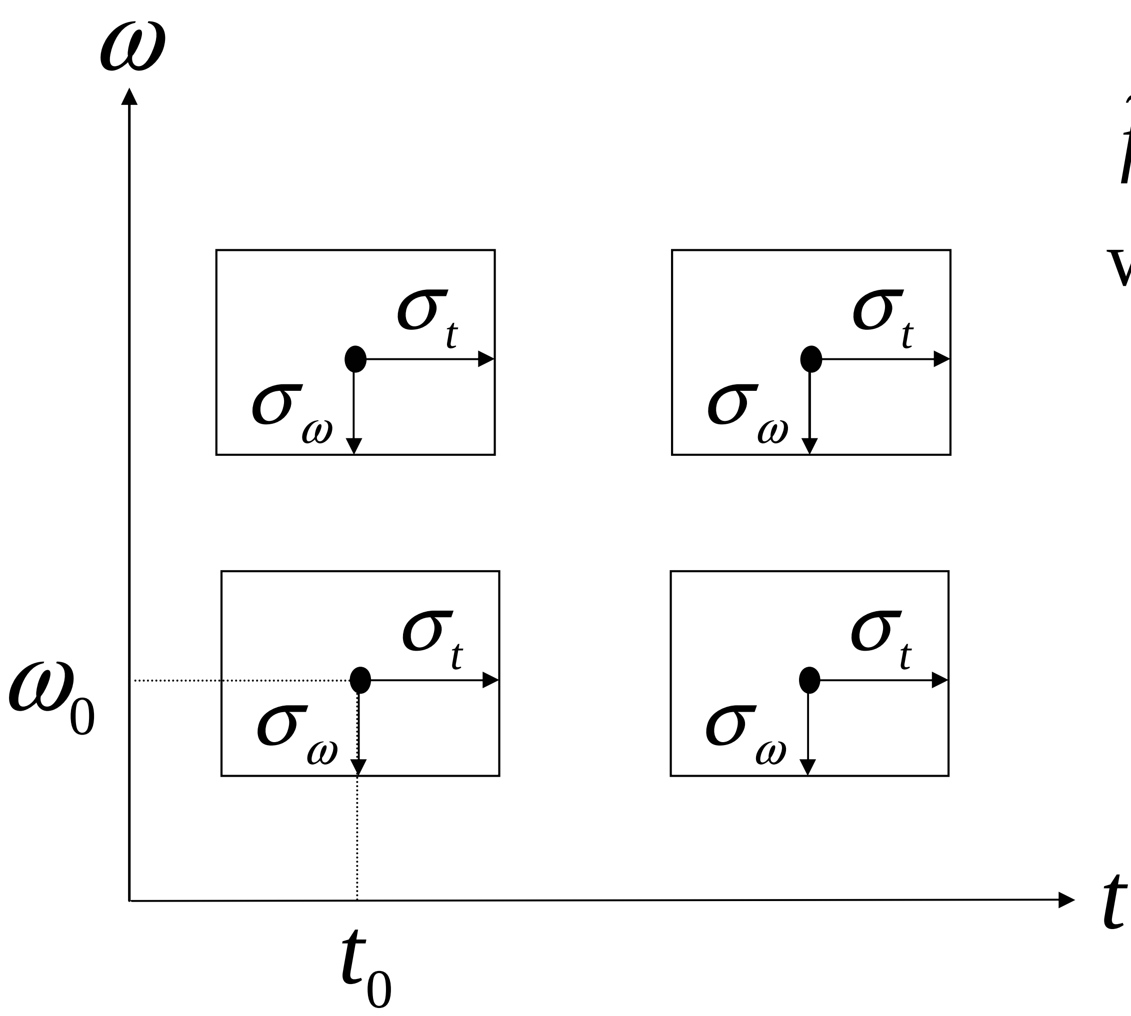

### Frequency



# Windowed Fourier Transform—Time Frequency Localization

## Time Frequency Localization




$\tilde{f}(\omega_0, t_0)$ describes $f$ within a resolution cell:

$$[t_0 \pm \sigma_t][\omega_0 \pm \sigma_\omega]$$

Uncertainty principle:

$$\sigma_t \sigma_\omega > \frac{1}{4\pi}$$

# Windowed Fourier—Reconstruction

## Reconstruction Formula

$$f(u) = \frac{1}{C} \iint g_{\omega,t}(u)\, \tilde{f}(\omega,t)\, d\omega dt \qquad C = \|g\|^2$$

$$f_t(u) = f(u)\overline{g}(u-t)$$

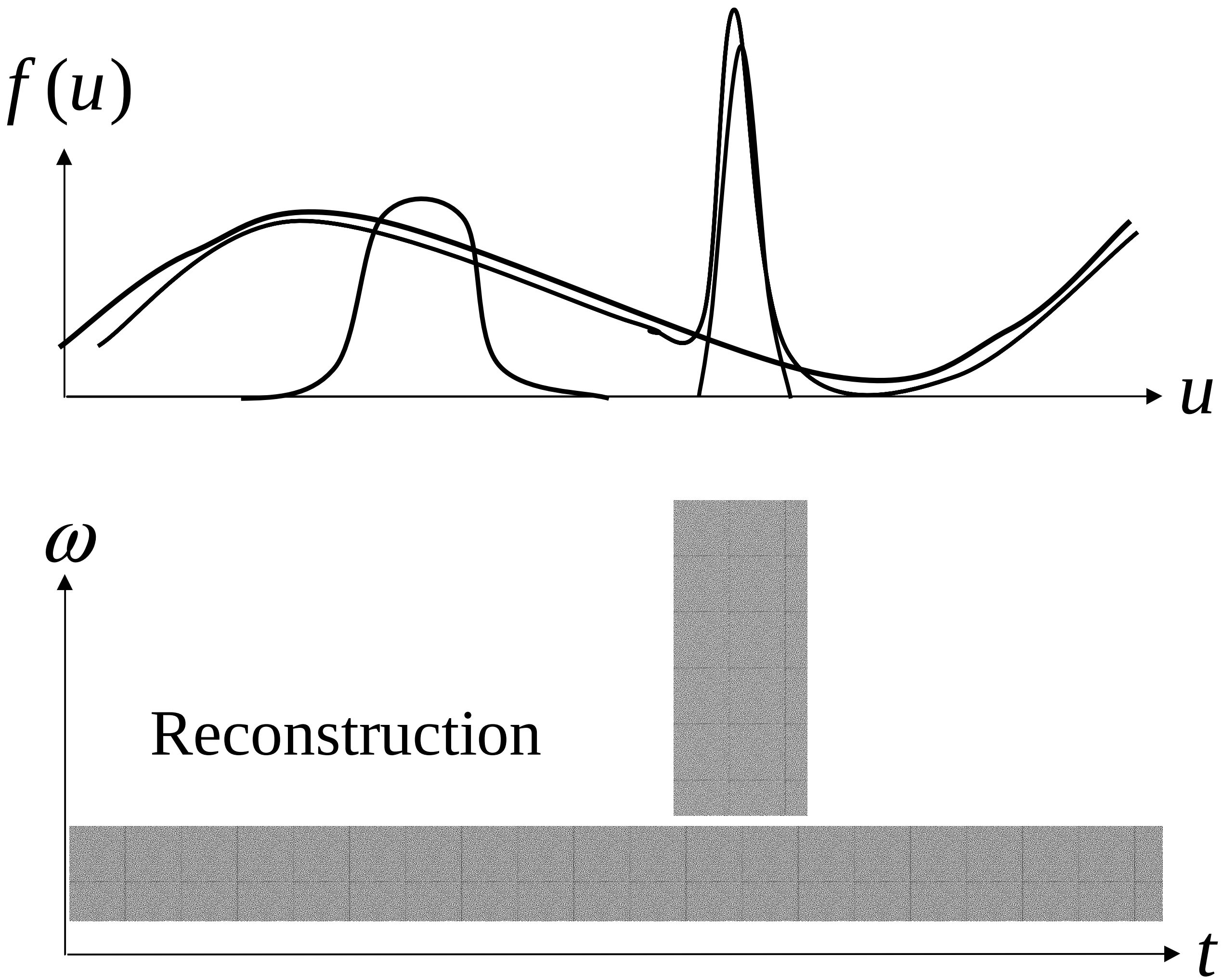
$$\tilde{f}(\omega,t) = \int_{-\infty}^{\infty} f_t(u) e^{-j2\pi\omega u} du$$

$$\overline{g}(u-t) f(u) = \int_{-\infty}^{\infty} \tilde{f}(\omega,t) e^{j2\pi\omega u} d\omega$$

$$f(u) \underbrace{\int_{-\infty}^{\infty} |g(u-t)|^2 dt}_{C} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{f}(\omega,t) \underbrace{g(u-t) e^{j2\pi\omega u}}_{g_{\omega,t}(u)} d\omega dt$$

# Windowed Fourier—Reconstruction

## Reconstruction Formula



# Windowed Fourier—Redundancy

## Redundancy

$$f(u)=\frac{1}{C}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\tilde{f}(\omega,t)\underbrace{g(u-t)e^{j2\pi\omega u}}_{g_{\omega,t}(u)}d\omega dt$$

Parseval for WFT

$$\tilde{f}(\omega,t)=\left\langle g_{\omega,t}(u),f(u)\right\rangle_{L^2(R)}=\frac{1}{C}\left\langle \tilde{g}_{\omega,t}(\omega',t'),\tilde{f}(\omega',t')\right\rangle_{L^2(R^2)}$$

$$\tilde{g}_{\omega,t}(\omega',t')=\left\langle g_{\omega,t}(u),g_{\omega',t'}(u)\right\rangle=K_g(\omega,t|\omega',t')=\int_{-\infty}^{\infty}\overline{g}_{\omega,t}(u)g_{\omega',t'}(u)du$$

$$\tilde{f}(\omega,t)=\frac{1}{C}\iint\underbrace{\overline{K}_g(\omega,t|\omega',t')}_{K_g(\omega',t'|\omega,t)}\tilde{f}(\omega',t')dt'd\omega'$$

-Values of $\tilde{f}$ are correlated

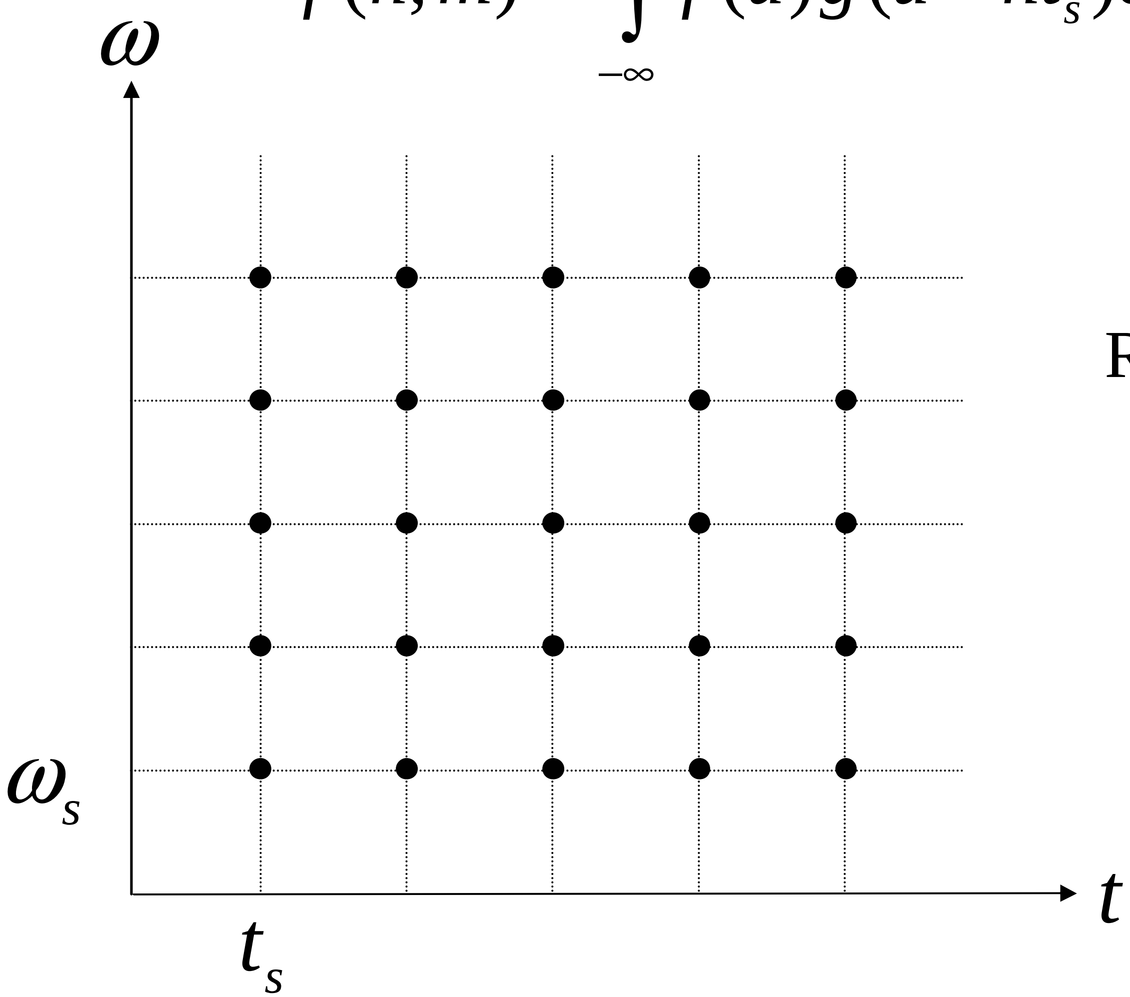
-Not any function $\tilde{f}(\omega,t)\in L^2$ can be a WFT

-$\tilde{f}(\omega,t)\in\mathcal{F},\mathcal{F}\subset L^2$ is a reproducing kernel Hilbert space

# Windowed Fourier Transform—Sampling Theorem

## Sampling Theorem

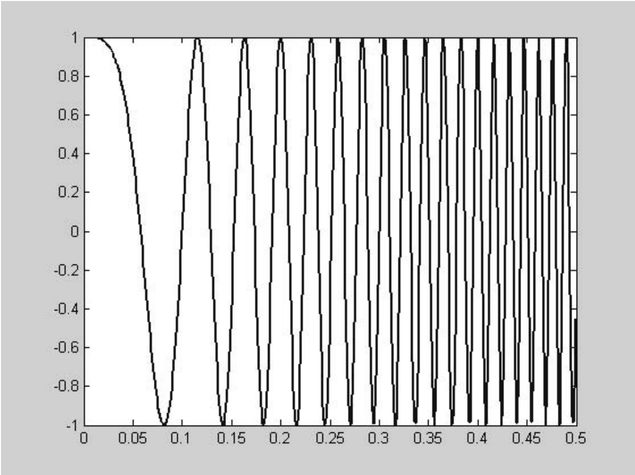
$$\tilde{f}(n,m) = \int_{-\infty}^{\infty} f(u)\bar{g}(u - nt_s)e^{-j2\pi m\omega_s u}du$$



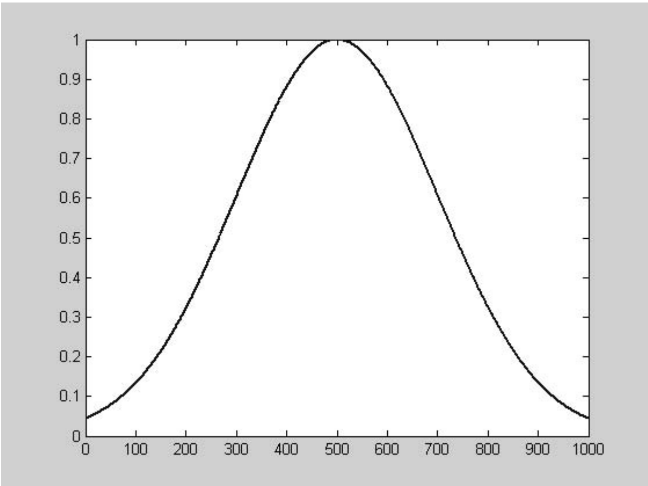

Reconstruction if

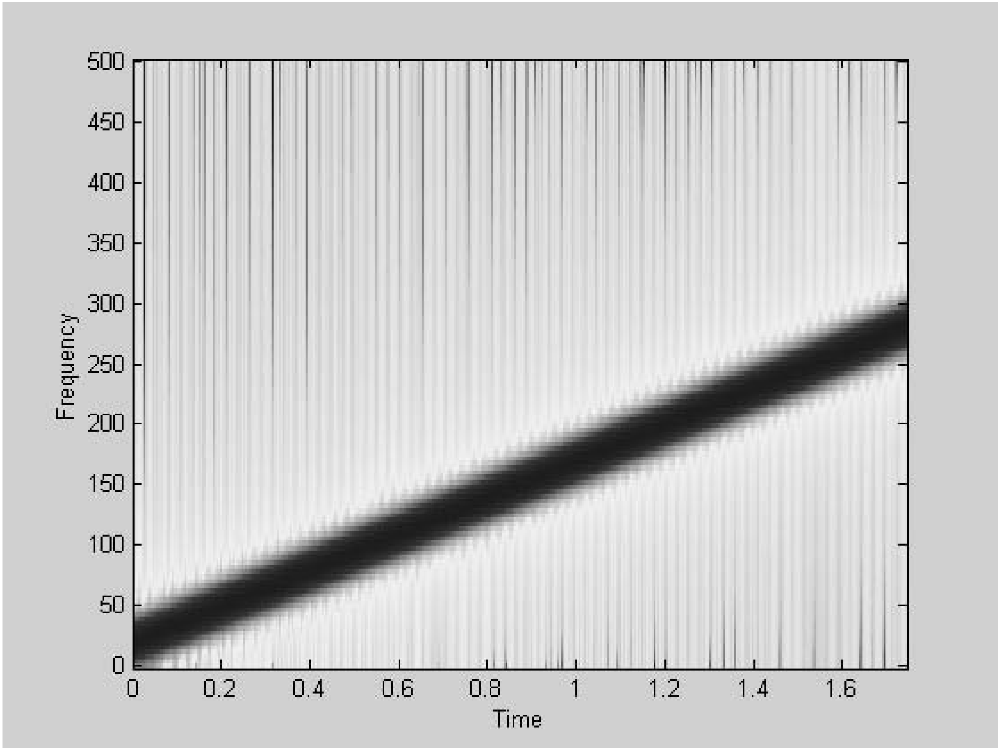
$$\omega_s t_s < 1$$

# Windowed Fourier Transform—Examples

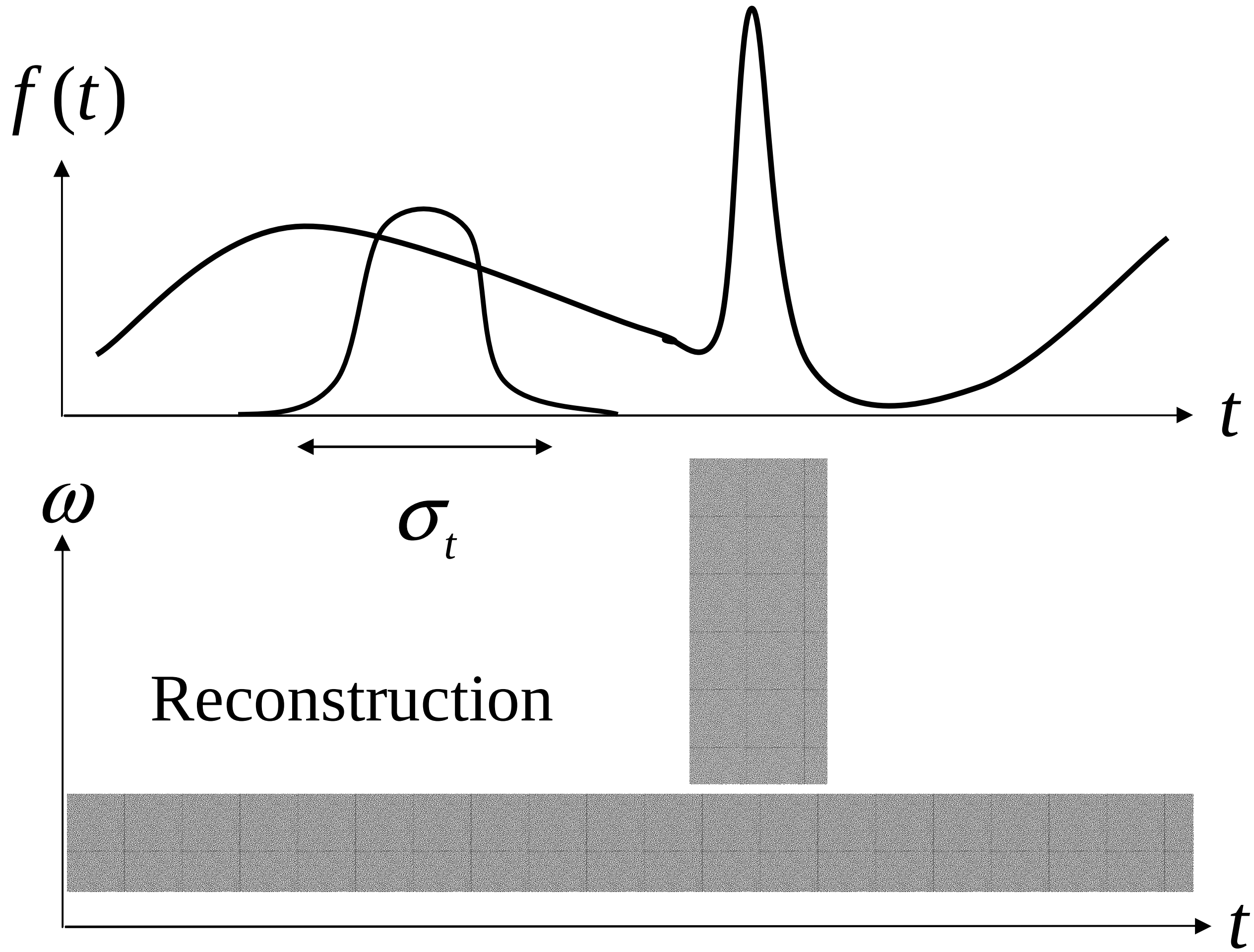



$$\cos(\pi u^2), \omega_{\text{inst}} = u$$



# Time Scale Analysis—Motivation




Features with time scales much shorter and longer than $\sigma_t$ have to be synthesized with many notes.

# Time Scale Analysis—Continuous Wavelet Transform

$$\psi_{s,t}(u) = \frac{1}{|s|^p}\ \psi\left(\frac{u-t}{s}\right), \quad \psi(u) \in L^2, s \neq 0, p > 0$$

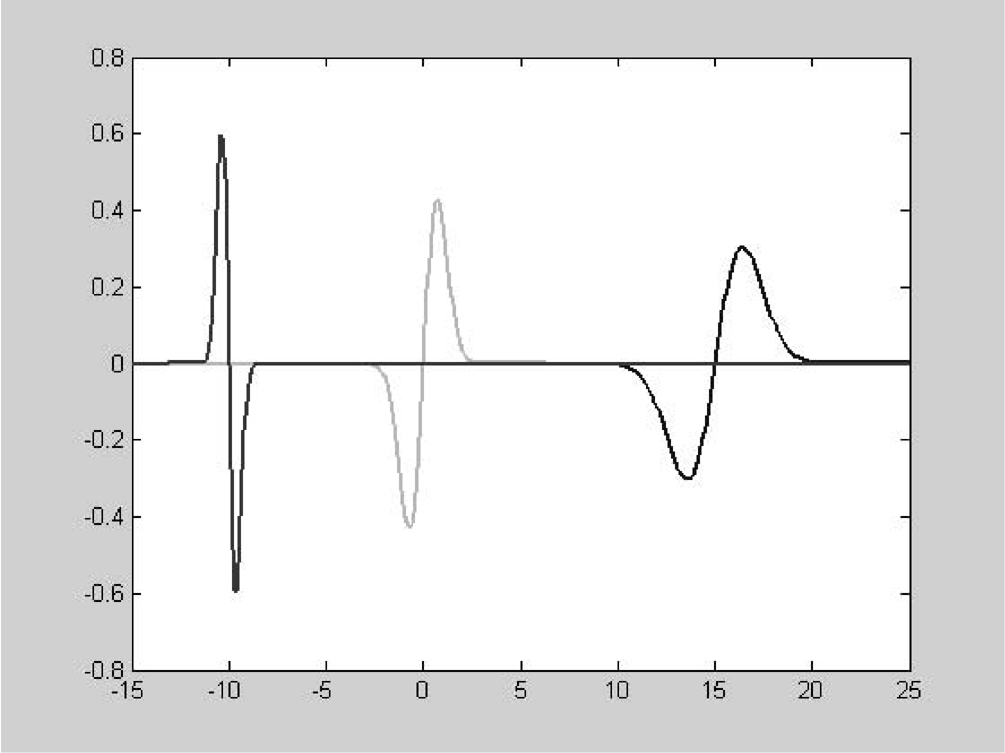
$$\tilde{f}(s,t) = \int_{-\infty}^{\infty} \bar{\psi}_{s,t}(u)\, f(u)du = \langle \psi_{s,t}, f \rangle, \quad f \in L^2$$

$\psi(u) = ue^{-u^2}$

$\psi_{-0.5,-10}$ red

$\psi_{1,0}$ green

$\psi_{2,15}$ blue



# Wavelet Transform—Reconstruction

## Reconstruction Formula

Admissibility condition: $$0 < C_{\pm} = \int_0^{\infty} \frac{\left|\hat{\psi}(\pm\omega)\right|^2}{\omega} d\omega < \infty$$

$$f(u) = \int_0^{\infty} \int_{-\infty}^{\infty} \psi^{s,t}(u)\, \tilde{f}(s,t) s^{2p-3} dt ds$$

$\left\{\psi^{s,t}\right\}$ reciprocal wavelet family of $\left\{\psi_{s,t}\right\}$

if $C_{-} = C_{+} = \frac{C}{2}$ then $\psi_{s,t} = \frac{C}{2}\psi^{s,t}, \left\{\psi_{s,t}\right\}$ is self reciprocal:

$$f(u) = \frac{2}{C} \int_0^{\infty} \int_{-\infty}^{\infty} \psi_{s,t}(u)\, \tilde{f}(s,t) s^{2p-3} dt ds$$

# Wavelet Transform—Reconstruction

Admissibility condition:

$$0 < C_{\pm} = \int_0^{\infty} \frac{\left|\hat{\psi}(\pm\omega)\right|^2}{\omega} d\omega < \infty$$

$$\hat{\psi}(0) = 0 \quad \int_{-\infty}^{\infty} \psi(u) du = 0$$

Symmetry condition:

$$C_{-} = C_{+}$$

if $\psi$ is symmetric, $\hat{\psi}$ is symmetric

if $\psi$ is real-valued, $\hat{\psi}(-\omega) = \overline{\hat{\psi}}(\omega)$

# Wavelet Transform—Plancherel&Parseval

## Plancherel Formula

$$\left\langle \tilde{f}, \tilde{f} \right\rangle_{\mathcal{L}} = \frac{1}{C_+ + C_-} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |s|^{2p-3} \left| \tilde{f}(s,t) \right|^2 ds dt$$

## Parseval Identity

$$\left\langle f, g \right\rangle_{L^2(\mathbf{R})} = \left\langle \tilde{f}, \tilde{g} \right\rangle_{\mathcal{L}} \quad \forall\ f, g \in L^2(\mathbf{R})$$

$$\left\langle \tilde{f}, \tilde{g} \right\rangle_{\mathcal{L}} = \frac{1}{C_+ + C_-} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |s|^{2p-3}\ \overline{\tilde{f}}(s,t) \tilde{g}(s,t) ds dt$$

# Wavelet Transform—Time Frequency Symmetry

$$p = 0.5, s > 0: \quad \psi_{s,t}(u) = \frac{1}{\sqrt{s}} \; \psi\left(\frac{u-t}{s}\right)$$

$$\hat{\psi}_s(\omega) = \sqrt{s}\; e^{-j2\pi\omega t}\hat{\psi}(s\omega)$$

$$\tilde{f}(s,t) = \left\langle \psi_{s,t}, f \right\rangle = \left\langle \hat{\psi}_{s,t}, \hat{f} \right\rangle = \int_{-\infty}^{\infty} \overline{\hat{\psi}}_{s,t}(\omega)\, \hat{f}(\omega) du$$

$$\tilde{f}(s,t) = \sqrt{s} \int_{-\infty}^{\infty} \overline{\hat{\psi}}(s\omega)\, \hat{f}(\omega)\, e^{j2\pi\omega t} d\omega$$

$$\tilde{f}(s,t) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} \overline{\psi}\left(\frac{u-t}{s}\right) f(u) du$$

## Time Frequency Localization

$\psi(u)$ is centered at $t_0$ with $\sigma_t$, $\hat{\psi}(\omega)$ is centered at $\omega_0$ with $\sigma_\omega$

$$\tilde{f}(s,t') = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} \overline{\psi}\left(\frac{u-t'}{s}\right) f(u)du$$

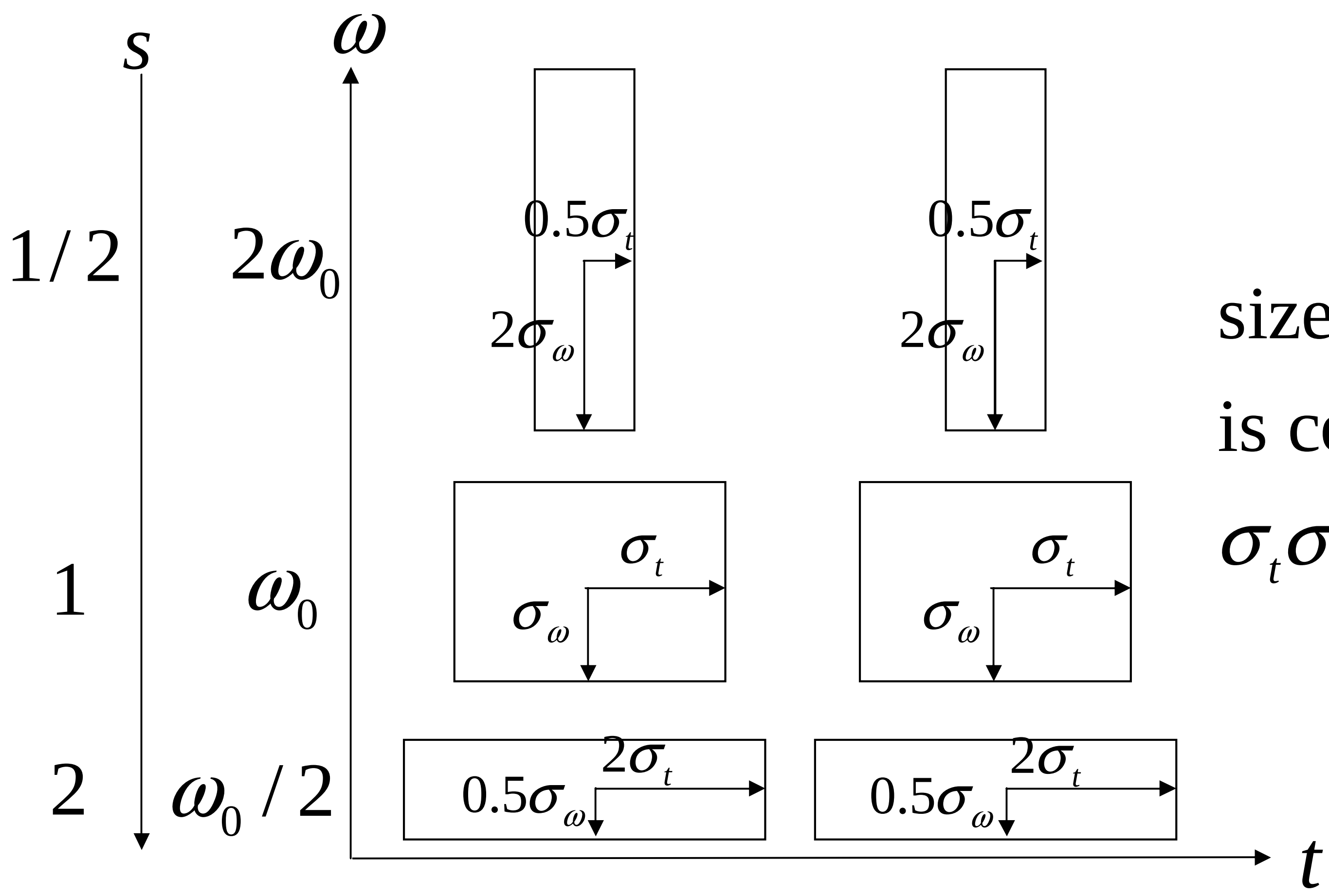
time window centered at $s\,t_0 + t'$ with $s\sigma_t$

$$\tilde{f}(s,t') = \sqrt{s} \int_{-\infty}^{\infty} \overline{\hat{\psi}}(s\omega)\, \hat{f}(\omega)\, e^{j2\pi\omega t'} d\omega$$

frequency window centered at $\omega_0 / s$ with $\sigma_\omega / s$

# Wavelet Transform—Time Frequency Localization

## Time Frequency Localization




size of resolution cells is constant: $\sigma_t \sigma_\omega$

# Wavelet Transform—Redundancy

## Redundancy

$$\left\langle f_1, f_2 \right\rangle_{L^2} = \left\langle \tilde{f}_1, \tilde{f}_2 \right\rangle_{\mathcal{L}} \qquad \tilde{f}(s,t) = \left\langle \psi_{s,t}, f \right\rangle_{L^2} = \left\langle \tilde{\psi}_{s,t}, \tilde{f} \right\rangle_{\mathcal{L}}$$

$$\tilde{\psi}_{s',t'} = \left\langle \psi_{s,t}, \psi_{s',t'} \right\rangle_{L^2} = K_{\psi}(s',t' \,|\, s,t)$$

$$\tilde{f}(s,t) = \iint |s|^{2p-3} K_{\psi}(s',t' \,|\, s,t)\, \tilde{f}(s',t')\, ds't'$$

-Values of $\tilde{f}$ are correlated

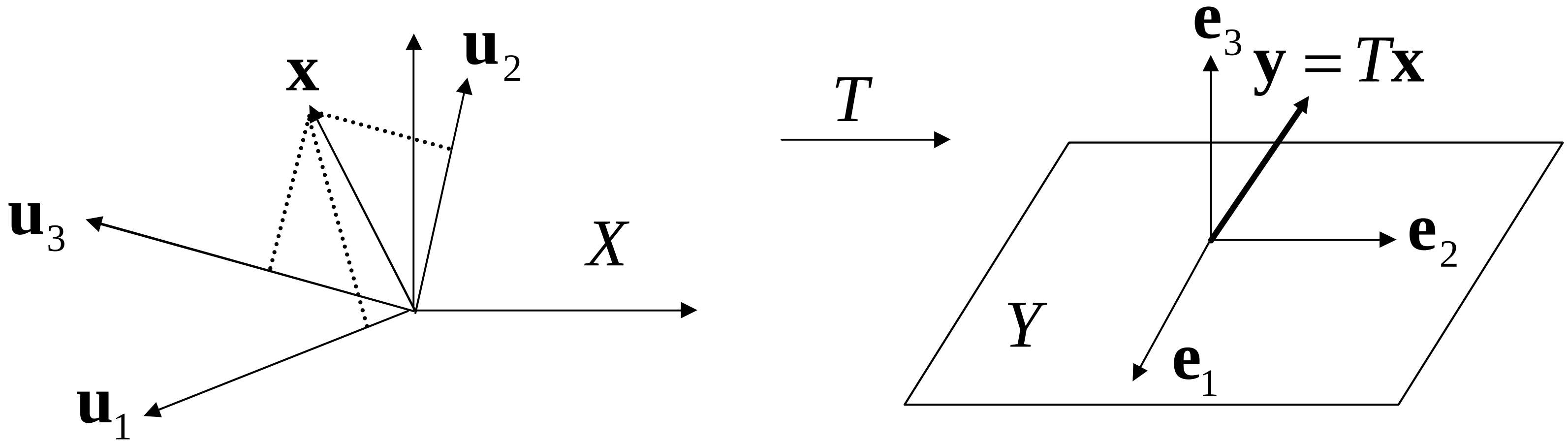
-Not any function $\tilde{f}(s,t) \in L^2$ can be a WFT

-$\tilde{f}(s,t) \in \mathcal{L}, \mathcal{L} \subset L^2$ is a reproducing kernel Hilbert space

# Wavelet Transform—Frames

## Notion of Frames

$$\tilde{f}(s,t) = \langle \psi_{s,t}, f \rangle$$




Hilbert space $X$, $\dim X = N$

$\{\mathbf{u}_1, \ldots, \mathbf{u}_M\} \in X$, $M > N$,

Hilbert space $Y$, $\dim Y = M$

$$T\mathbf{x} = \sum_j \langle \mathbf{x}, \mathbf{u}_j \rangle \mathbf{e}_j, \ T : X \to Y$$

$$\langle \mathbf{x}, T^* \mathbf{y} \rangle = \langle T\mathbf{x}, \mathbf{y} \rangle, T^* : Y \to X$$

Unlike the basis the frames $\{\mathbf{u}_1, \ldots, \mathbf{u}_M\}$ can be linearly dependent

# Wavelet Transform—Frames

if $\mathbf{x} \in X$ is uniquely determined by $T\mathbf{x} \in Y$

then $\{\mathbf{u}_1, ..., \mathbf{u}_M\} \in X$ is a frame of $X$

$T$ is injective

$\{\mathbf{u}_1, ..., \mathbf{u}_M\} \in X$ is a frame of $X$ if:

$$A\|\mathbf{x}\|^2 \leq \|T\mathbf{x}\|^2 \leq B\|\mathbf{x}\|^2 \quad \forall \mathbf{x} \in X, B \geq A \geq 0$$

Reconstruction of $\mathbf{x}$ from $\mathbf{y} = T\mathbf{x}$:

$$\mathbf{x} = S\mathbf{y} = (T^*T)^{-1}T^*\mathbf{y}$$

$G = T^*T$ is selfadjoint $\rightarrow$ eigenvalues are real

Frames are the framework for continuous and discrete wavelets

# Wavelet Transform—Discrete Wavelets

$$\tilde{f}(s,t) = \iint |s|^{2p-3} K_\psi(s',t' | s,t)\, \tilde{f}(s',t') ds't'$$

Redundancy: Discrete wavelets

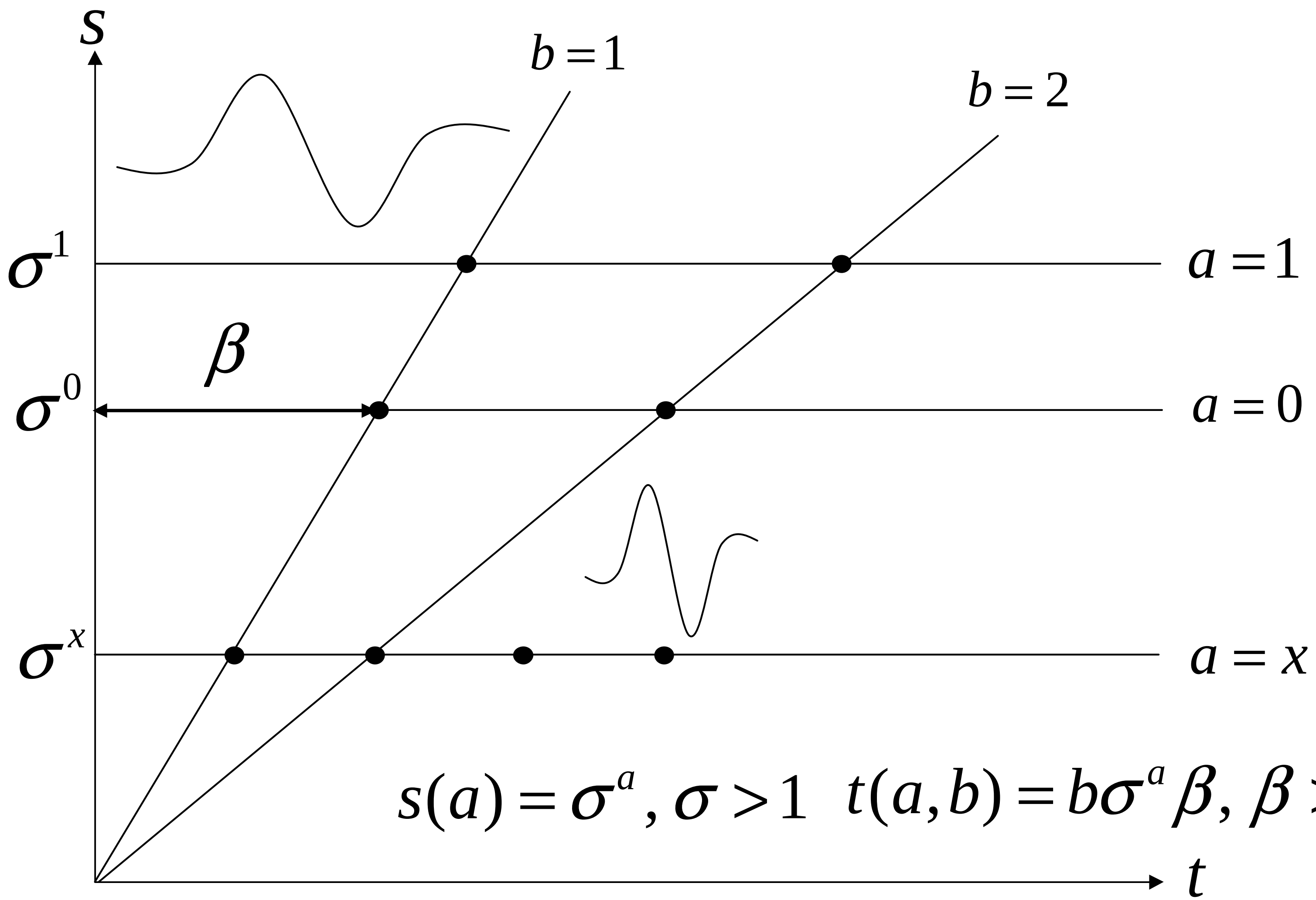
$$\psi(s,t),\ s,t \in \mathbf{R} \text{ to } \psi(a,b),\ a,b \in \mathbf{Z}$$

Exponential sampling

$$s(a) = \sigma^a,\ \sigma > 1, \text{ elementary dilation step}$$

$$t(a,b) = b\sigma^a \beta,\ \beta > 0$$

## Sampling of the $s,t$ plane




$$s(a) = \sigma^a, \sigma > 1 \quad t(a,b) = b\sigma^a\beta, \beta > 0$$

# Wavelet Transform—Multiresolution Analysis

## Multiresolution analysis of discrete signals

$$f(n), n \in \mathbf{Z}$$

Sampling $\tilde{f}(s,t)$ on a dyadic grid, orthogonal wavelets:

$$\sigma = 2, \beta = 1: s = 2^a, t = b2^a \quad a, b \in \mathbf{Z}$$

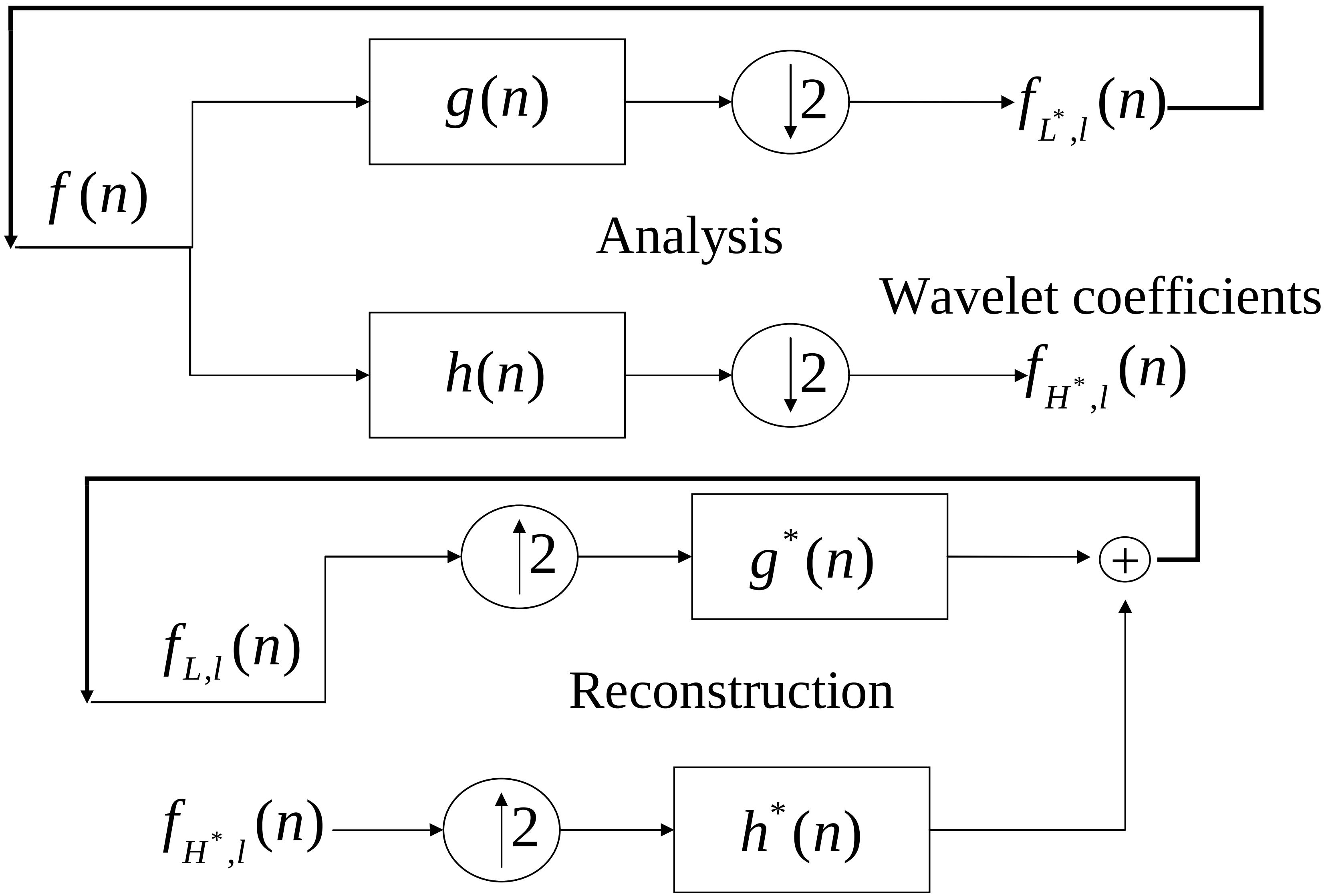
Two half band filters with impulse response $h(n), g(n)$

$$f_H(n) = f(n) * h(n) = \sum_k f(k)h(n-k)$$

$$f_L(n) = f(n) * g(n) = \sum_k f(k)g(n-k)$$

Downsampling: $f_{H^*}(n) = f_H(2n)$, $f_{L^*}(n) = f_L(2n)$

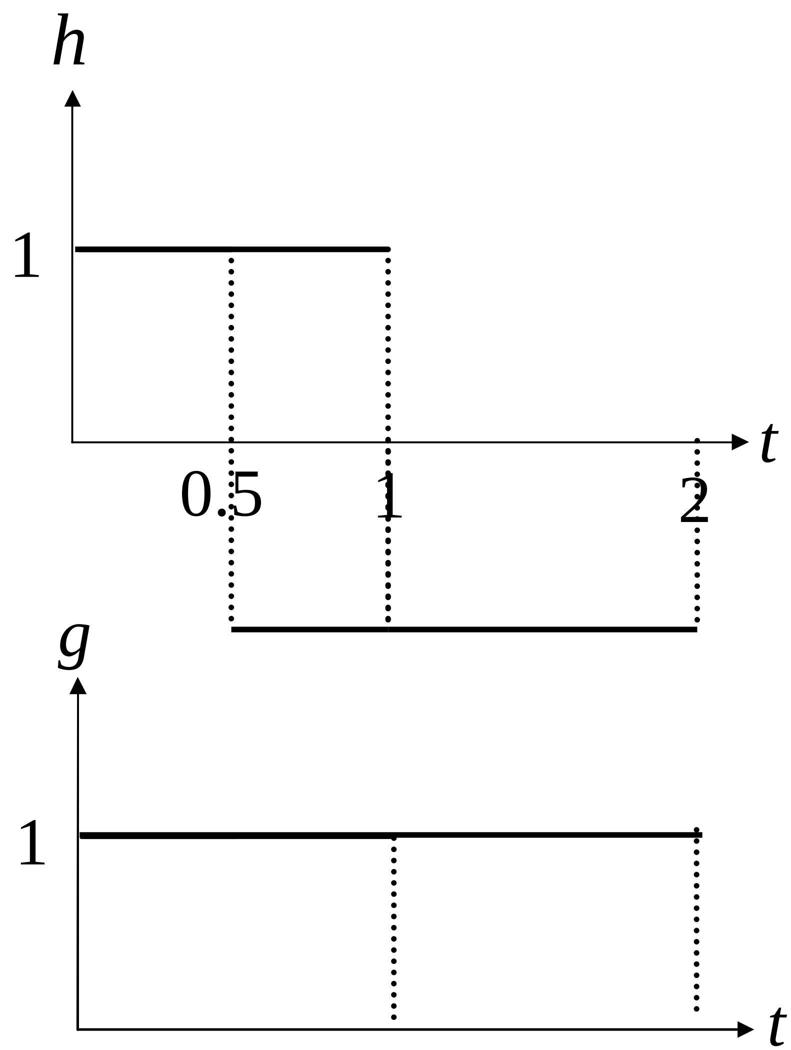
# Wavelet Transform—Multiresolution Analysis




$g(n), h(n)$ are a Quadrature mirror filter

# Wavelet Transform—Multiresolution Analysis

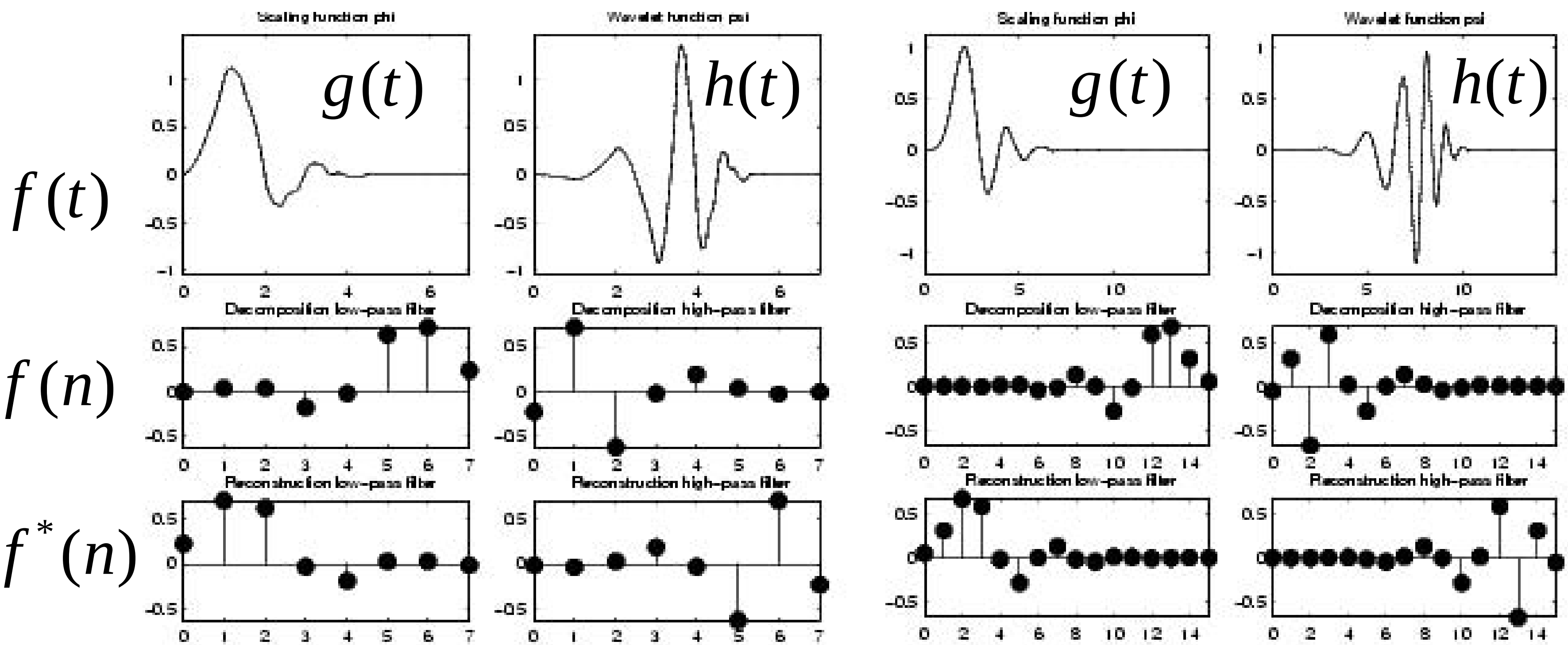
## Haar Wavelets




*h* are orthonormal
easy to compute but
poor frequency resolution

# Wavelet Transform—Multiresolution Analysis

## Daubechie Wavelets




*Matlab Documentation*

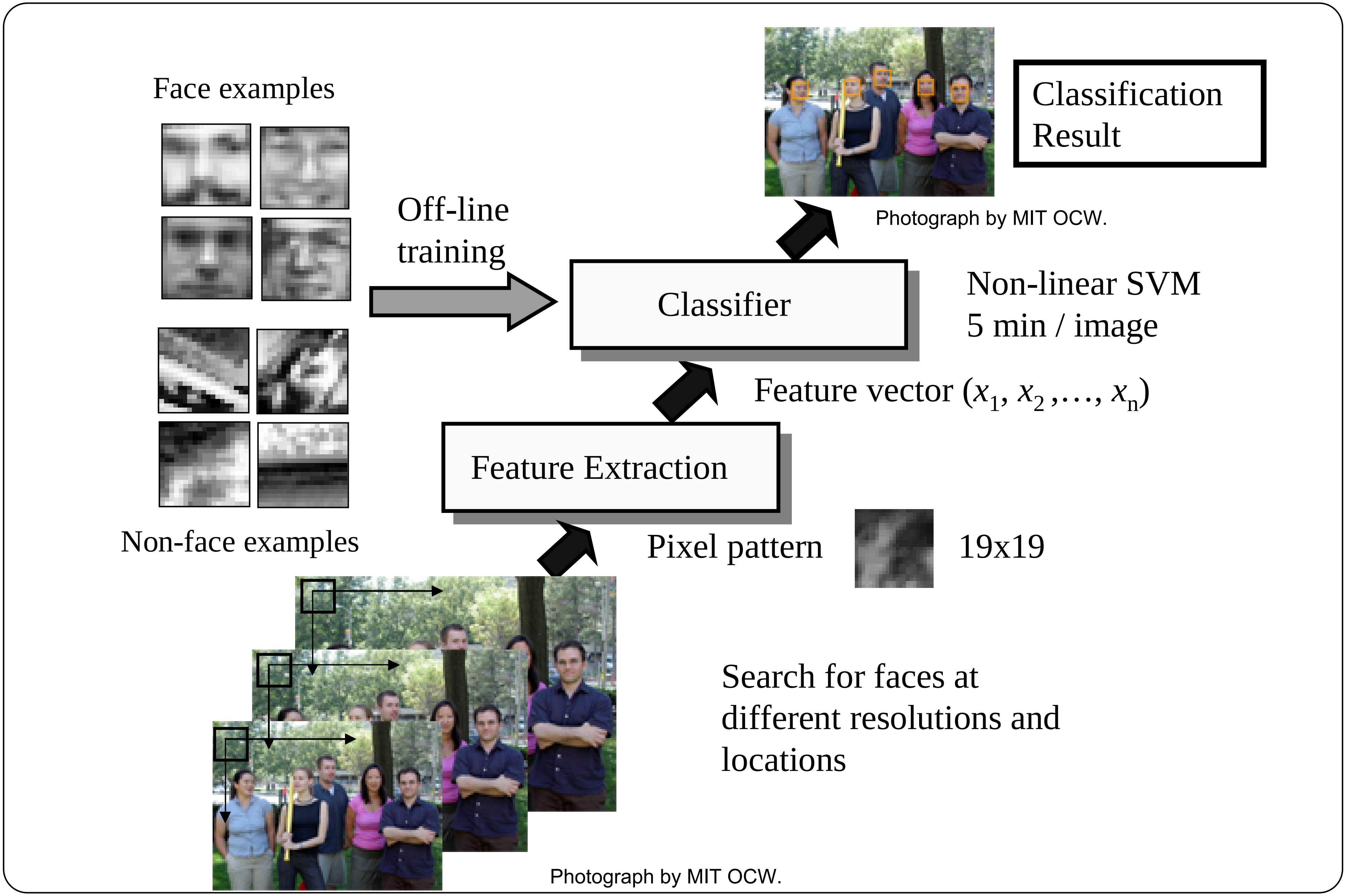
# Haar Wavelets (Matlab Toolbox)

Screenshot from Matlab Toolbox removed due to copyright reasons.

# Wavelet Transform—Applications

- Image Compression
- Texture Analysis
- De-noising
- Features for Object Detection and Recognition

# My Face Detector in 2000

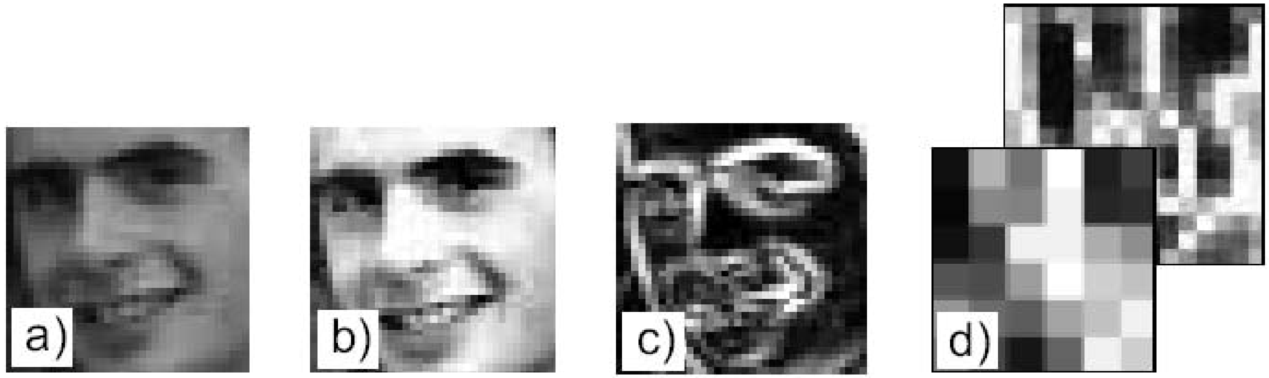



Photograph by MIT OCW.

Photograph by MIT OCW.

# My Face Detector in 2000

## My experiments with different types of features



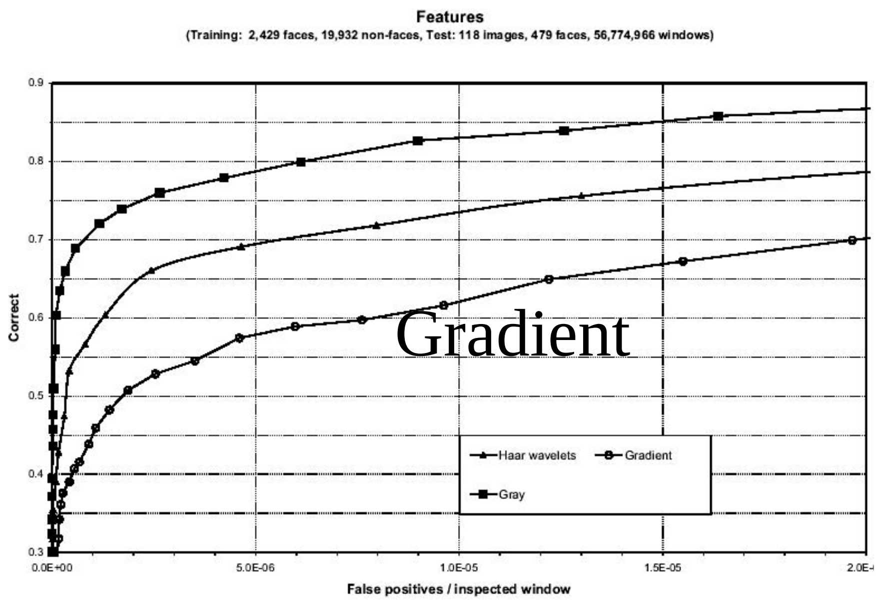

Original

Histogram equalized

Gradients

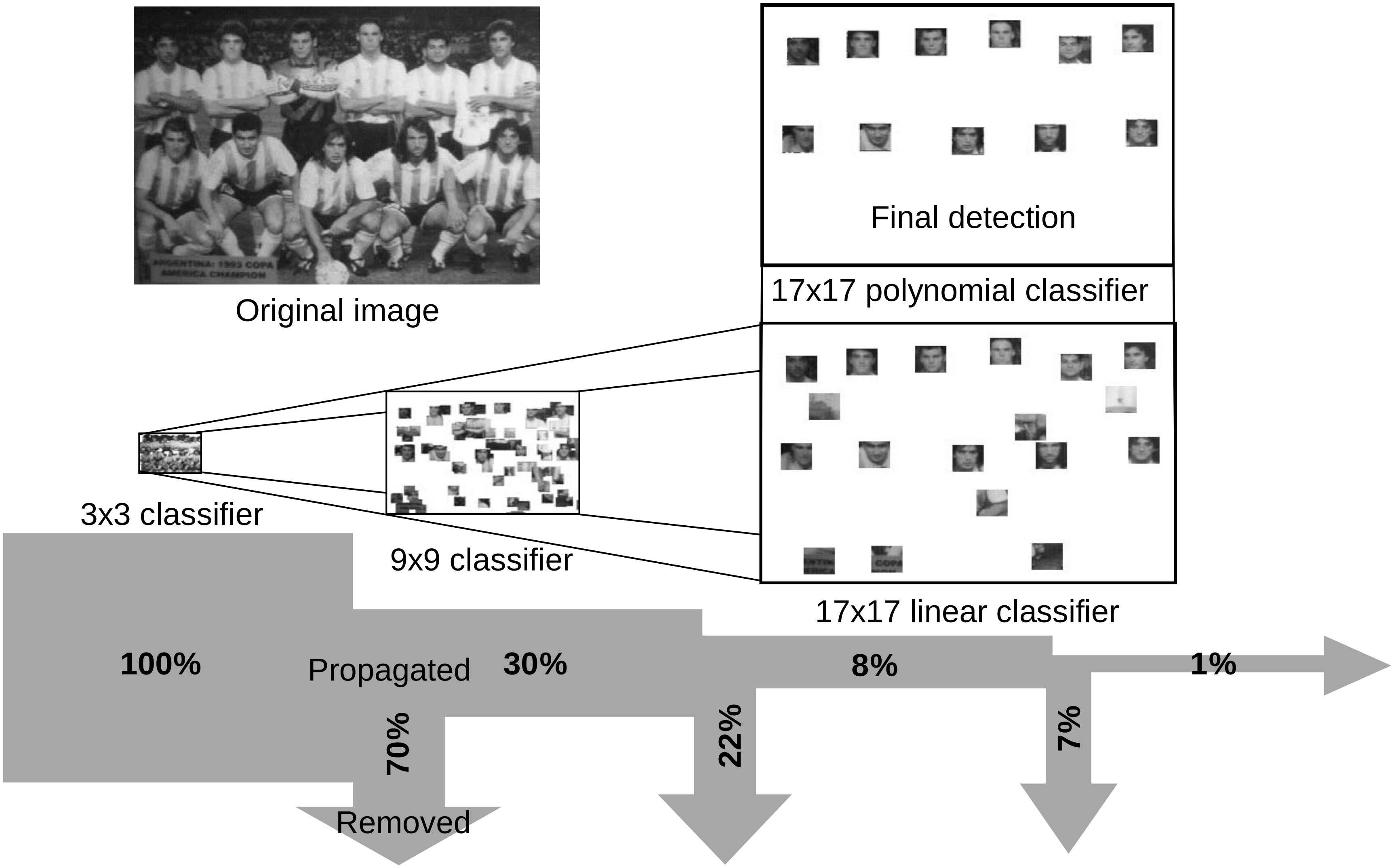
Haar Wavelets

Photographs courtesy of CMU/VASC Image Database at http://vasc.ri.cmu.edu/idb/html/face/frontal_images/




*AI Memo 2000*

# Speeding-up Face Detection




Photograph courtesy of CMU/VASC Image Database at http://vasc.ri.cmu.edu/idb/html/face/frontal_images/

# Hierarchical Face Detection—Heisele, CVPR 2001

| System | Typical detection time | Speed-up factor |
|---|---|---|
| Single $2^{nd}$ degree polynomial SVM | 271 s | – |
| Single $2^{nd}$ degree polynomial SVM + Feature reduction | 63.8 s | 4.25 |
| 3-Level hierarchy + Feature reduction | 1.6 s | 170 |

**Table 1. Computing time for a 320×240 image processed on a dual Pentium III with 733 MHz. The original image was rescaled in 5 steps to detect faces at resolutions between 26×26 and 60×60 pixels.**

# Viola&Jones Detector—Viola&Jones, CVPR 2001

Speed is proportional to the average number of features computed per sub-window.

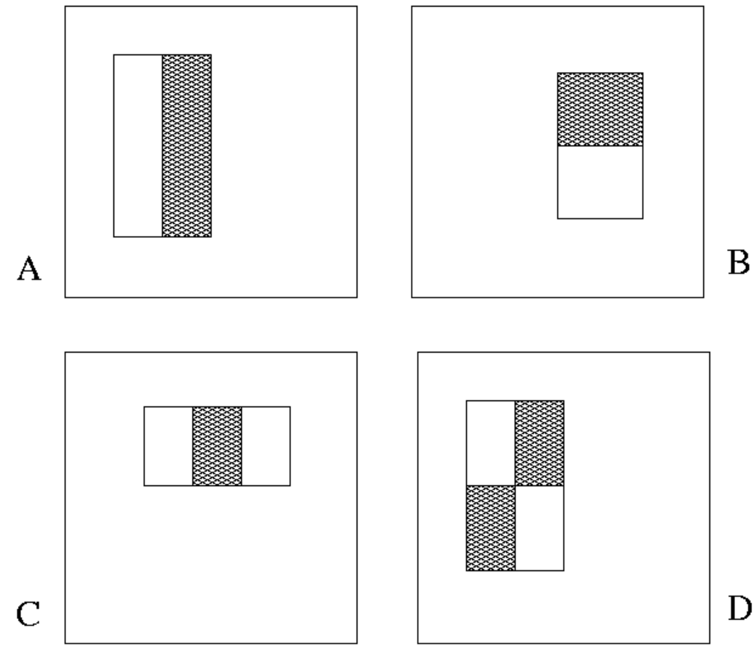
On the MIT+CMU test set, an average of 9 features out of a total of 6061 are computed per sub-window.

**On a 700 Mhz Pentium III, a 384x288 pixel image takes about 0.067 seconds to process (15 fps).**

Roughly 15 times faster than Rowley-Baluja-Kanade and 600 times faster than Schneiderman-Kanade.

# Viola&Jones Detector—Image Features

Photo removed due to copyright considerations.




"Rectangle filters"

**Similar to Haar wavelets**

Differences between sums of pixels in adjacent rectangles

$$h_t(x) = \begin{cases} +1 & \text{if } f_t(x) > \theta_t \\ -1 & \text{otherwise} \end{cases}$$

$$160{,}000 \times 100 = 16{,}000{,}000$$

Unique Features

# Viola&Jones Detector

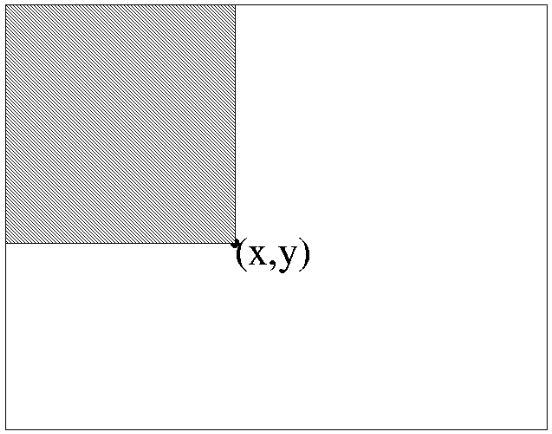
How exactly does it work??
Read the CVPR 2001 paper…….
or wait until the lecture on object detection by Mike Jones

# Integral Image—Viola&Jones CVPR 2001

Define the Integral Image

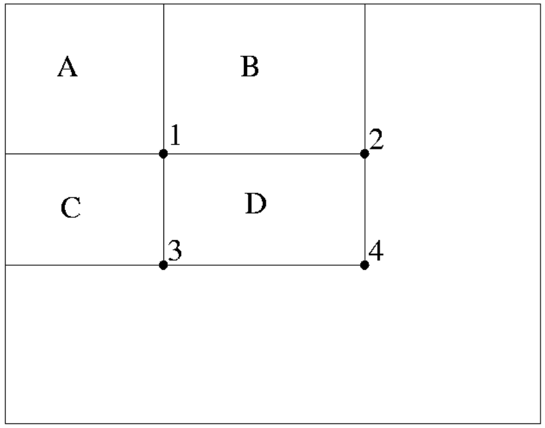
Any rectangular sum can be computed in constant time:

$$I'(x,y) = \sum_{\substack{x' \le x \\ y' \le y}} I(x',y')$$




Rectangle features can be computed as differences between rectangles

$$\begin{aligned} D &= 1 + 4 - (2 + 3) \\ &= A + (A + B + C + D) - (A + C + A + B) \\ &= D \end{aligned}$$




Series of photos and figures from Rapid object detection using a boosted cascade of simple features. Viola, P., and M. Jones. Computer Vision and Pattern Recognition, 2001. CVPR 2001. *Proceedings of the 2001 IEEE Computer Society Conference* 1 (2001): I-511 - I-518. Graphs courtesy of IEEE. 

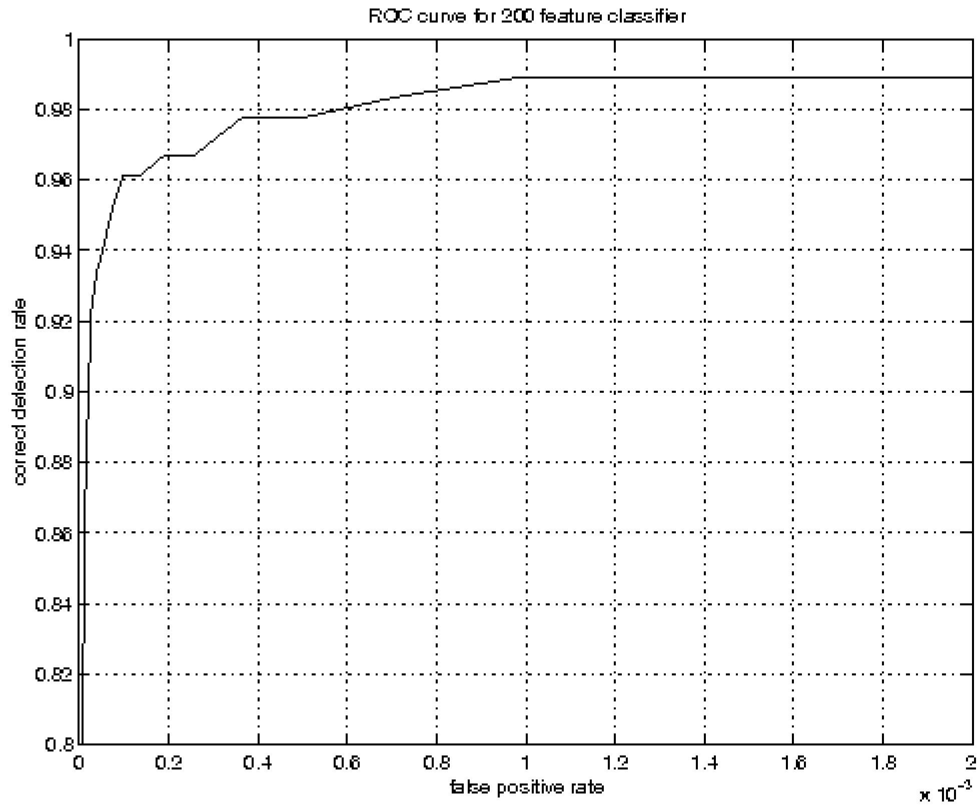
# Example Classifier for Face Detection—Viola&Jones CVPR 2001

A classifier with 200 rectangle features was learned using AdaBoost

95% correct detection on test set with 1 in 14084
false positives.

Not quite competitive...

Photographs removed due to
copyright considerations.




ROC curve for 200 feature classifier

Series of photos and figures from Rapid object detection using a boosted cascade of simple features. Viola, P., and M. Jones. Computer Vision and Pattern Recognition, 2001. CVPR 2001. *Proceedings of the 2001 IEEE Computer Society Conference* 1 (2001): I-511 - I-518. Graphs courtesy of IEEE. Copyright 2001 IEEE. Used with Permission.

# Literature

Ingrid Daubechies "Ten Lectures on Wavelets"
*For mathematicians only, many proofs*

Gerald Kaiser "A Friendly Guide to Wavelets"
*Not friendly, but easier to understand than above*

Christian Blatter "Wavelets—A Primer"
*more friendly than Kaiser…*

Stephane Mallat "Multifrequency Channel Decomposition"
IEEE Acoustics Speech & Signal Proc. 1989
*One of the early papers on multiresolution analysis*

# Homework

- Some proofs (optional)

- Template matching with Fourier transform
  Shape representation with Fourier descriptors
  (stick to instructions)

- Playing with wavelets