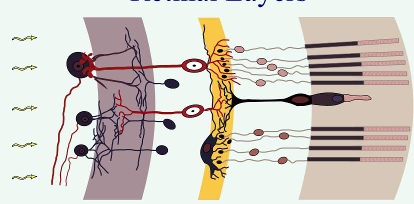
6. Perception

Outline

A. Biology of Perception
B. Primary Visual Cortex
C. Object Recognition and "What" Pathway
D. Attention and "How" Pathway

A. Biology of Perception

Retinal Layers

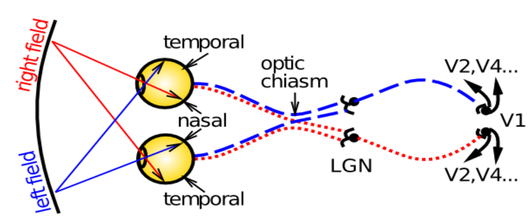


From right:
(1) rods and cones (tan),
(2) bipolar and horizontal cells (yellow),
(3) amacrine and retinal ganglion cells (purple)

Retinal Cells

- Retina is CNS tissue
- Nonuniform distribution of cells
  - rods in periphery specialized for sensitivity and motion
  - cones in macula/fovea specialized for color & form (acuity)
  - humans are *foveating* animals
- Information compression:
  - 120 million rods (low light)
    - huge convergence of rods on bipolars ⇒ sensitivity
  - 6 million cones (color and high acuity)
  - 1 million RGC (retinal ganglion cells)
    - estimated to transmit 10^9 bits/sec
    - spontaneously active; information conveyed by change in rate

Optic Pathway



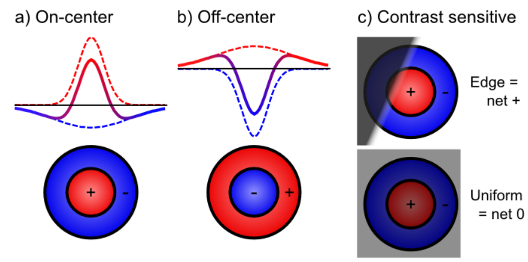
(slide < O'Reilly)

Key Organizing Principles

- Transduction of different information
  - — wavelength (rods; blue, green, red cones)
  - — spatial frequency (resolution)
  - — motion
- Topographic organization
  - — contrasting similar information
- Filtering to extract relevant information

2/21/17 COSC 494/594 CCN 7

Retinal Contrast Filtering



2/21/17 COSC 494/594 CCN 8

(slide < O'Reilly)

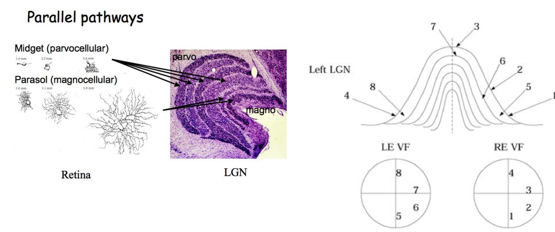
LGN of the Thalamus

- A "relay station," but also much more
- Organizes different types of information into different layers with aligned retinotopic maps
- Performs dynamic processing: magnocellular motion processing cells, attentional processing
- On- and off-center information from retina is preserved in LGN

2/21/17 COSC 494/594 CCN 9

(slide based on Frank)

Structure of LGN

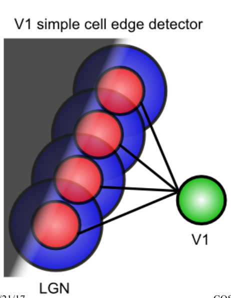


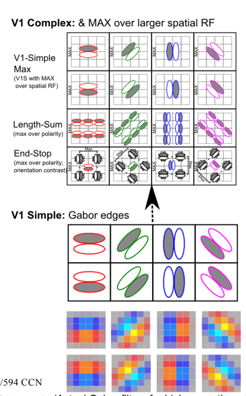
- Cells have monocular input
- Six layers alternate input from two eyes (RGC)

2/21/17 COSC 494/594 CCN 10

(figs < Heeger)

V1

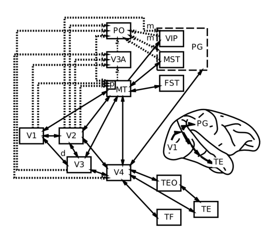




2/21/17 COSC 494/594 CCN 11

(slide < O'Reilly) (Actual Gabor filters for high-res pathway)

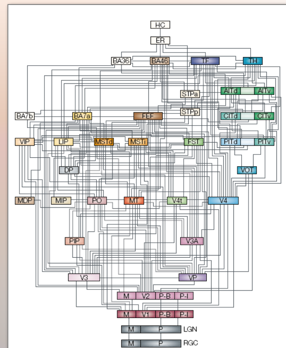
"What" vs. "Where" Pathways



- "What" ignores differences in location, illumination, size, rotation
- "Where" emphasizes location, size, and ignores object identity

2/21/17 COSC 494/594 CCN 12

Hierarchy of Macaque Visual Areas



2/21/17 COSC 494/594 CCN 13

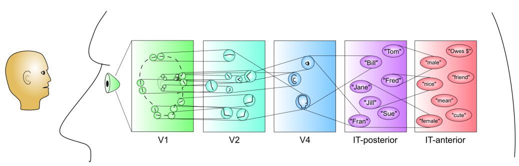
(fig. from Van Essen & al. 1992)

Principal Regions in "What" Pathway

- V1: Primary Visual Cortex
  - — encodes image in terms of oriented edges
- V2: Secondary Visual Cortex
  - — encodes in terms of intersections & junctions
- V4
  - — more complex features over wider range of locations
- PIT: Posterior Inferotemporal (IT) Cortex
  - — location & size invariant object recognition
  - — includes FFA (fusiform face area)
- AIT: Anterior IT Cortex
  - — abstract/semantic visual information

2/21/17 COSC 494/594 CCN 14

Hierarchy of Visual Detectors



2/21/17 COSC 494/594 CCN 15

B. Primary Visual Cortex

What is the origin of detectors for oriented bars of light?

2/21/17 COSC 494/594 CCN 16

Self-Organization of V1 Orientation Selective Neurons

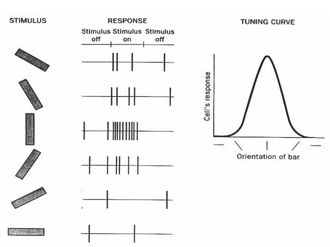


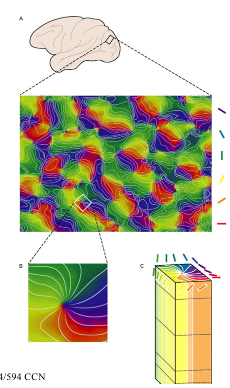
FIGURE 4.8 Response of a single cortical cell to bars presented at various orientations.

2/21/17 COSC 494/594 CCN 17

(slide < O'Reilly)

Topographic Maps

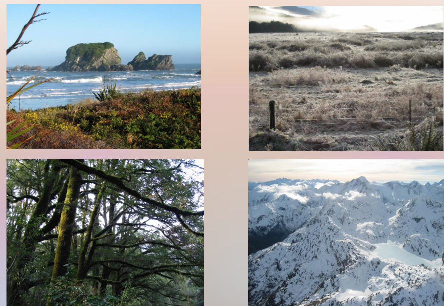
- Map of orientations
- Hypercolumn: Full set of coding for each position
- Pinwheel can arise from learning and lateral connectivity: not hard-wired!



2/21/17 COSC 494/594 CCN 18

(slide based on O'Reilly)

What is Common?

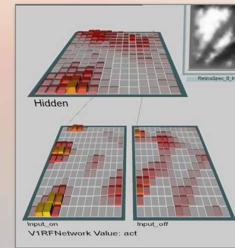


2/21/17 COSC 494/594 CCN 19

(slide < O'Reilly)

V1Rf: Simulating One Hypercolumn

- Natural visual scenes are preprocessed by passing them (separately) through layers of on-center and off-center inputs
- Hidden layer: edge detectors seen in layers 2/3 of V1; Layer 4 (input) just represents unoriented on/off inputs like LGN (but can be modulated by attention)
- Circular neighborhood of lateral excitatory connectivity in Hidden layer



2/21/17 COSC 494/594 CCN 20

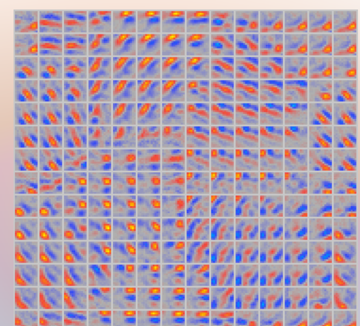
(slide based on Frank)

emergent demonstration: V1Rf

2/21/17 COSC 494/594 CCN 21

Self-Organized Topography

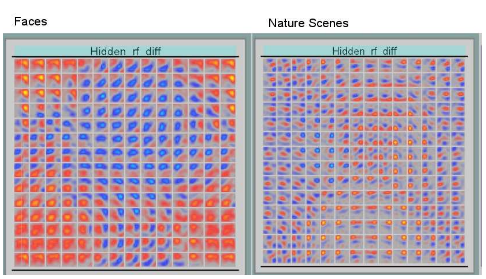
Model shows how documented V1 properties can result from interactions between learning, architecture (connectivity), and structure of environment



2/21/17 COSC 494/594 CCN 22

(slide < O'Reilly)

Faces vs. Natural Scenes



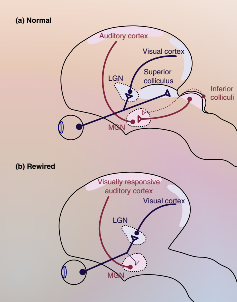
Some differences, but pinwheels still emerge

2/21/17 COSC 494/594 CCN 23

(slide < Frank)

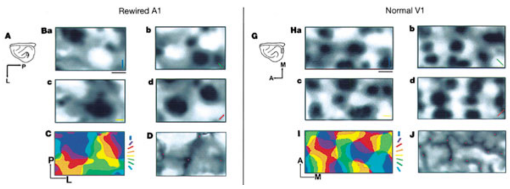
Rewiring Cortex

- Experiments by Mriganka Sur & colleagues (MIT)
- What happens if retinal axons are redirected into auditory thalamus (MGN) instead of its usual inputs?
- Answer: Auditory cortex (A1) develops orientation columns and retinotopic maps similar to V1
- Animals experience activity in A1 as visual perception



2/21/17 COSC 494/594 CCN 24

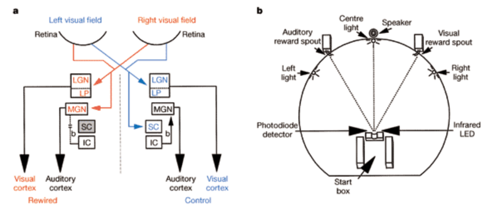
Orientation Columns in A1



Orientation columns develop in A1 similar to those in V1

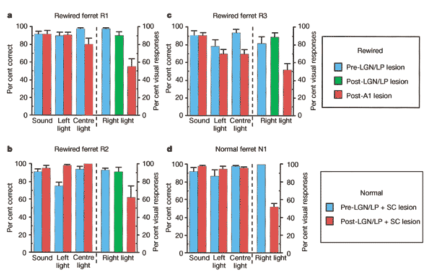
2/21/17 COSC 494/594 CCN 25

Are They Having Visual Experiences?



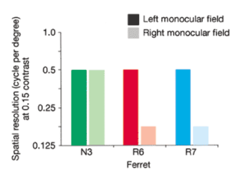
2/21/17 COSC 494/594 CCN 26

They Are Having Visual Experiences



2/21/17 COSC 494/594 CCN 27

Visual Acuity



- There is less visual acuity in the rewired pathway
- Suggests there may be intrinsic factors in organization of auditory cortex as well as extrinsic factors

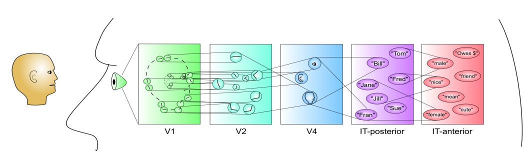
2/21/17 COSC 494/594 CCN 28

C. Object Recognition and "What" Pathway

How do we recognize objects (across locations, sizes, rotations with wildly different retinal images)?

2/21/17 COSC 494/594 CCN 29

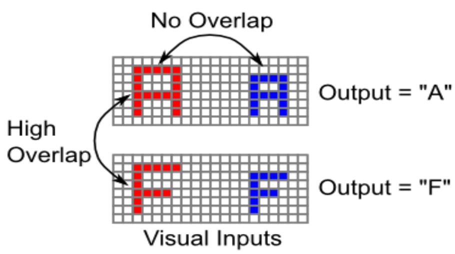
Invariant Object Recognition



2/21/17 COSC 494/594 CCN 30

(slide < O'Reilly)

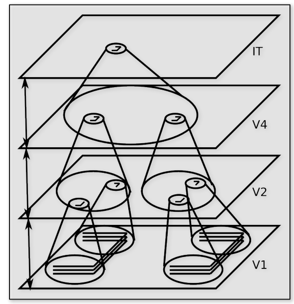
It's Hard



2/21/17 COSC 494/594 CCN 31

(slide < O'Reilly)

Invariant Object Recognition

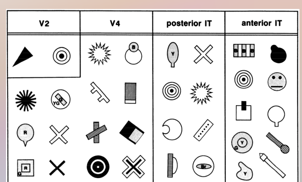


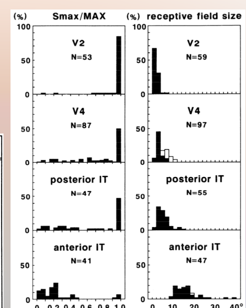
- Hierarchy of increasing:
  - Feature complexity
  - Spatial invariance
- Increasing RF size:
  - Conjunction of features (to form more complex objects)
  - Collapsing over location information ("spatial invariance")
- Strong match to RF's in corresponding brain areas

2/21/17 COSC 494/594 CCN 32

(slide based on O'Reilly & Frank)

Biological Data: Increasing Complexity and Invariance

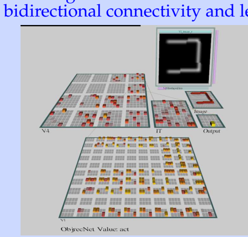




2/21/17 COSC 494/594 CCN 33

(slide < O'Reilly)

The Model: combining Fukushima with convolutional neural nets, bidirectional connectivity and learning!



V1 = oriented line (edge) detectors, hard-coded
V4 units encode conjunctions of V1 edges across a subset of space
Each IT unit pays attention to all of V4

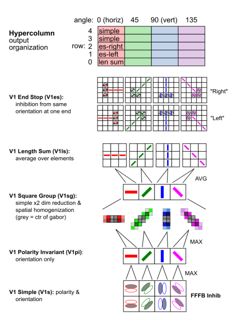
(V2 omitted here, important for figure-ground etc)

2/21/17 COSC 494/594 CCN 34

(slide < Frank)

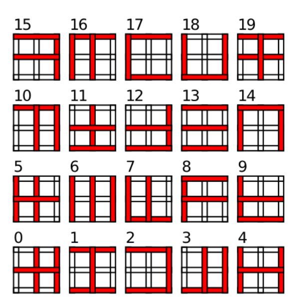
V1 Receptive Fields

- 4×5 hypercolumns
- Two rows of simple cells at 4 orientations and two polarities
- Two rows of end-stop complex cells
- One row of length-sum complex cells
- 50% overlap with adjacent hypercolumns



2/21/17 COSC 494/594 CCN 35

Simple Textbook Test



2/21/17 COSC 494/594 CCN 36

(slide < O'Reilly)

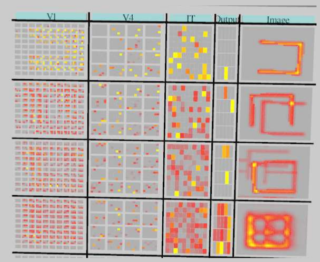
Activation-Based Receptive Fields

- How do we plot receptive fields for V4?
- Receiving weights show which V1 units a V4 unit responds to, but they don't show what thing in the world the unit responds to
- Solution: Show the network lots of input patterns.
- Then, display a composite of all the input patterns that activate the unit (weighted by activity).

2/21/17 COSC 494/594 CCN 37

(slide < Frank)

V4 Receptive Fields



- Some V4 units code for location-specific conjunctions of V1 features
- This will show up as a sharp receptive field for Image input

2/21/17 COSC 494/594 CCN 38

(slide < Frank)

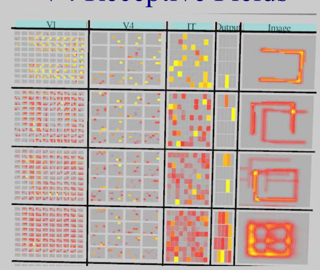
V4 Receptive Fields

- Some V4 units code for simple features in a location invariant way
- This will show up as smeary parallel lines in Image input

2/21/17 COSC 494/594 CCN 39

(slide < Frank)

V4 Receptive Fields



- Can also look at which Output units tend to get active for any given V4 unit
- Generally a given V4 unit is associated with multiple objects

2/21/17 COSC 494/594 CCN 40

(slide < Frank)

3D Object Recognition Test

3D models from Google SketchUp

100 categories

9–10 objects per category

2 objects left out for testing

+/– 20° horiz depth rotation + 180° flip

0–30° vertical depth rotation

14° 2D planar rotations

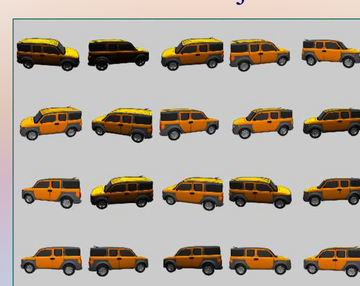
25% scaling

30% planar translations

2/21/17 COSC 494/594 CCN 41

http://grey.colorado.edu/CompCogNeuro/index.php/CU3D (slide < O'Reilly)

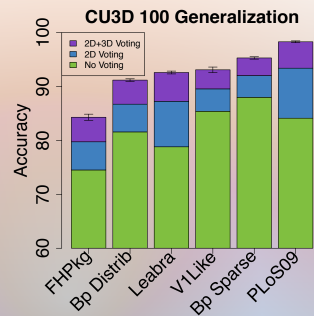
Depth & Lighting Variations for One Object



2/21/17 COSC 494/594 CCN 42

(slide < Frank)

Generalization to Novel Category Exemplars – Better than 90%



2/21/17 COSC 494/594 CCN 43

(slide < O'Reilly)

emergent demonstration: Objrec

2/21/17 COSC 494/594 CCN 44

D. Attention and "How" Pathway

Why is visual system split into what/where pathways?

Why does parietal damage cause attention problems (neglect)?

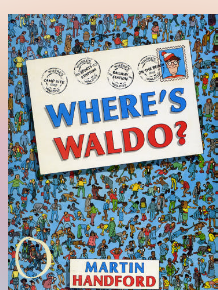
2/21/17 COSC 494/594 CCN 45

Some Functions of Dorsal Pathway

- "Where" pathway (spatial relations)
  - — visual attention (this chapter)
- But more broadly "how" pathway
  - — maps perception to action (next chapter)
- Numerical and mathematical processing
- Representation of abstract relationships
- Modulation of episodic memory
- Aspects of executive control

2/21/17 COSC 494/594 CCN 46

Spatial Attention and Neglect

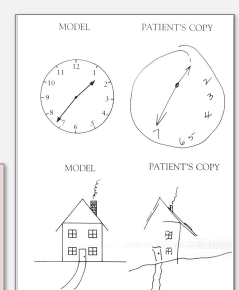


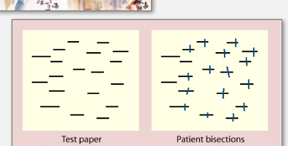
2/21/17 COSC 494/594 CCN 47

(slide < O'Reilly)

Hemispatial Neglect

Mainly from injuries to right parietal cortex



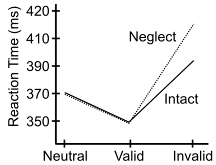


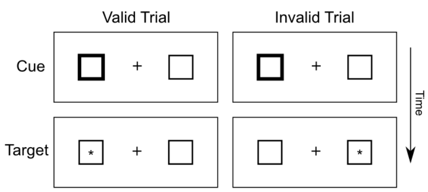
2/21/17 COSC 494/594 CCN 48

(slide based on O'Reilly)

Posner Task

- Valid cues speed performance (relative to "no cue" condition)
- Invalid cues slow performance (relative to "no cue" condition)



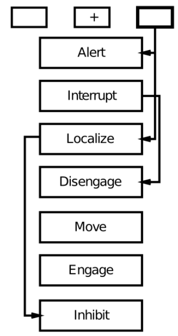


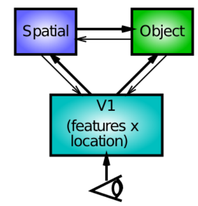
- Patients perform normally in the "neutral" (no cue) condition, regardless of where the target is presented
- Patients benefit just as much as controls from valid cues
- Patients are hurt more than controls by invalid cues

2/21/17 COSC 494/594 CCN 49

(slide based on O'Reilly)

Models: Boxology vs. Biology





2/21/17 COSC 494/594 CCN 50

(slide < O'Reilly)

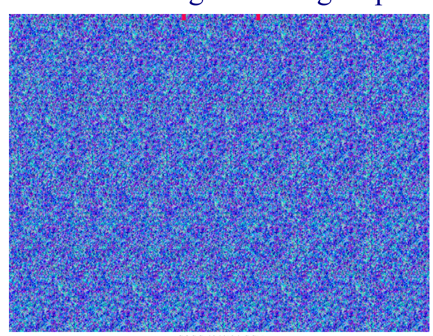
Posner Task Simulation

- Model explains the basic finding that valid cues speed target processing, while invalid cues hurt
- Also explains finding that patients with small unilateral parietal lesions benefit normally from valid cues in ipsilateral field but are disproportionately hurt by invalid cues
- No need to posit "disengage" module
- Also explains finding of **neglect** of contralateral visual field after large, unilateral parietal lesions when some stimulus is present in ipsilateral field ("extinction")

2/21/17 COSC 494/594 CCN 51

(slide < Frank)

What's Missing? Lacking Depth..



2/21/17 COSC 494/594 CCN 52

(slide < O'Reilly)

Supplementary

2/21/17 COSC 494/594 CCN 53

Gabor Uncertainty Principle and Gabor Elementary Functions

MacLennan, B. J. *Gabor Representations of Spatiotemporal Visual Images*. University of Tennessee, Knoxville, Computer Science Department technical report CS-91-144, September 1991

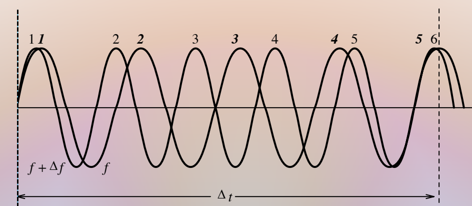
2/22/17 COSC 494/594 CCN 54

Dennis Gabor

- Dennis Gabor (1900–79) is the father of holography (1947, 1971 Nobel Prize in Physics)
- "the future cannot be predicted, but futures can be invented"
- Developed a theory of information (1946) complementary to Shannon's theory
- Gabor Uncertainty Principle based on same mathematics as derivation of Heisenberg Uncertainty Principle
- Nearly optimal Gabor representations are used in primary visual cortex

2/22/17 COSC 494/594 CCN 55

Time to Detect Difference in Frequency

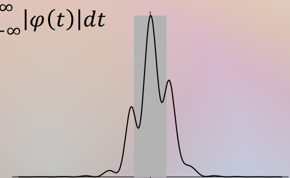


$$\Delta f \Delta t \geq 1$$

2/21/17 COSC 494/594 CCN 56

A Little More Formally…

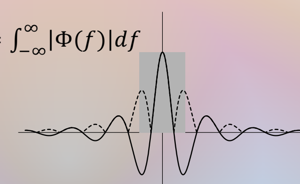
- **Nominal duration** (Δt) = duration of rectangular pulse with same area as signal and height equal to amplitude at origin
- Hence, $\Delta t|\varphi(0)| = \int_{-\infty}^{\infty}|\varphi(t)|dt$
- Some details omitted



2/21/17 COSC 494/594 CCN 57

A Little More Formally (2)

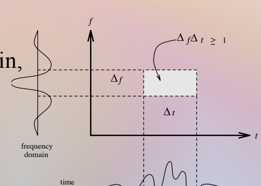
- **Nominal bandwidth** (Δf) of spectrum = width of rectangular pulse with height equal to spectrum's amplitude at origin and same area as absolute value of spectrum
- Hence, $\Delta f|\Phi(0)| = \int_{-\infty}^{\infty}|\Phi(f)|df$



2/21/17 COSC 494/594 CCN 58

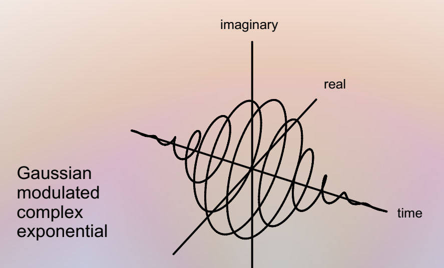
A Little More Formally (3)

- Computing the Fourier transform at origin, $|\Phi(0)| \leq \Delta t|\varphi(0)|$
- So $\Delta t \geq |\Phi(0)|/|\varphi(0)|$
- Computing the inverse Fourier transform at origin, $|\varphi(0)| \leq \Delta f|\Phi(0)|$
- So $\Delta f \geq |\varphi(0)|/|\Phi(0)|$
- Hence, $\Delta f \Delta t \geq 1$



2/21/17 COSC 494/594 CCN 59

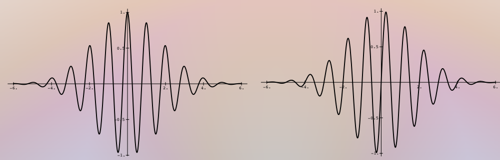
1D Gabor Elementary Function



2/21/17 COSC 494/594 CCN 60

Orthogonal Components of 1D Gabor

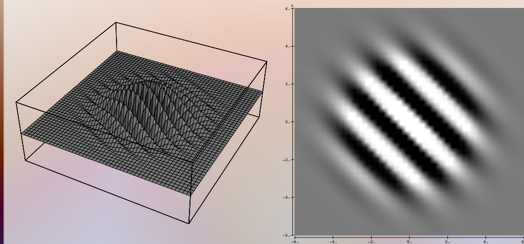
Cosine (real) **Sine (imaginary)**



$$e^{ift} = \cos ft + i \sin ft$$

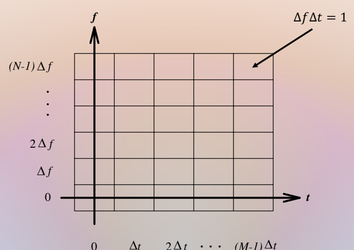
2/21/17 COSC 494/594 CCN 61

Real Part of 2D Gabor Elem. Function



2/21/17 COSC 494/594 CCN 62

Maximum Logon Content of Signal



2/21/17 COSC 494/594 CCN 63

Maximum Logon Content

- If $T = M\Delta t$ is the duration and $F = N\Delta f$ is the bandwidth
- The maximum number of logons *MN* is achieved when $\Delta t \Delta f = 1$ (i.e., Gabor elementary functions)
- In general, the area doesn't have to be divided into rectangles of the same shape, so long as area is 1
- So the maximum logon content is *TF* (duration times bandwidth)
- Any such signal can be represented uniquely as a sum of *TF* Gabor elementary functions

2/22/17 COSC 494/594 CCN 64

Gabor Representations

- Any "finite energy" function ψ of finite duration *D* and finite bandwidth *F* is equal to a linear superposition of Gabor elementary functions:

$$\psi(t) = \sum_{j=0}^{M-1} \sum_{k=0}^{N-1} a_{jk} C_{jk}(t) + b_{jk} S_{jk}(t)$$

where $C_{jk}(t) = e^{-\pi(t-j\Delta t)^2/\alpha^2} \cos[2\pi k \Delta f(t - j\Delta t)]$

and $S_{jk}(t) = e^{-\pi(t-j\Delta t)^2/\alpha^2} \sin[2\pi k \Delta f(t - j\Delta t)]$

- The same applies in higher dimensions.

2/22/17 COSC 494/594 CCN 65

Gabor Filters in Early Vision

- Measurements of receptive fields of simple cells in cat visual cortex have show them to be like Gaussian-modulated sinusoids (Jones & Palmer, 1987)
- Daugman (1984, 1985, 1993) showed 97% of them are statistically indistinguishable from the odd- or even-symmetric parts of a 2D Gabor elementary function
- Adjacent simple cells have grating patches that are 90° out of phase, but matched in preferred orientation and frequency
- And more… (MacLennan, 1991)

2/22/17 COSC 494/594 CCN 66