"Real World Applications: Unsupervised Learning

(**Duda Hart**):**Single Pixel Blind Sources Separation**"

- for Real-World Variant Mixing Matrix [A(x,y,t)]in the remote sensing & earlier cancer detection

- Machine IQ—Current Status of Computational Intelligence (CI) needs out of box creativity taking multidisciplinary CI approach---

- NSF Brain Sciences $20M per annum Initiative circa 2008-2018

Harold Szu,

a Founder (INNS), Fellows (IEEE,OSA,SPIE,AIMBE), Academician (RAS) ***szu_h@yahoo.com***

Dir. & Res. Prof., Digital Media RF Lab, Dept. ECE, GWU ***szuh@gwu.edu***

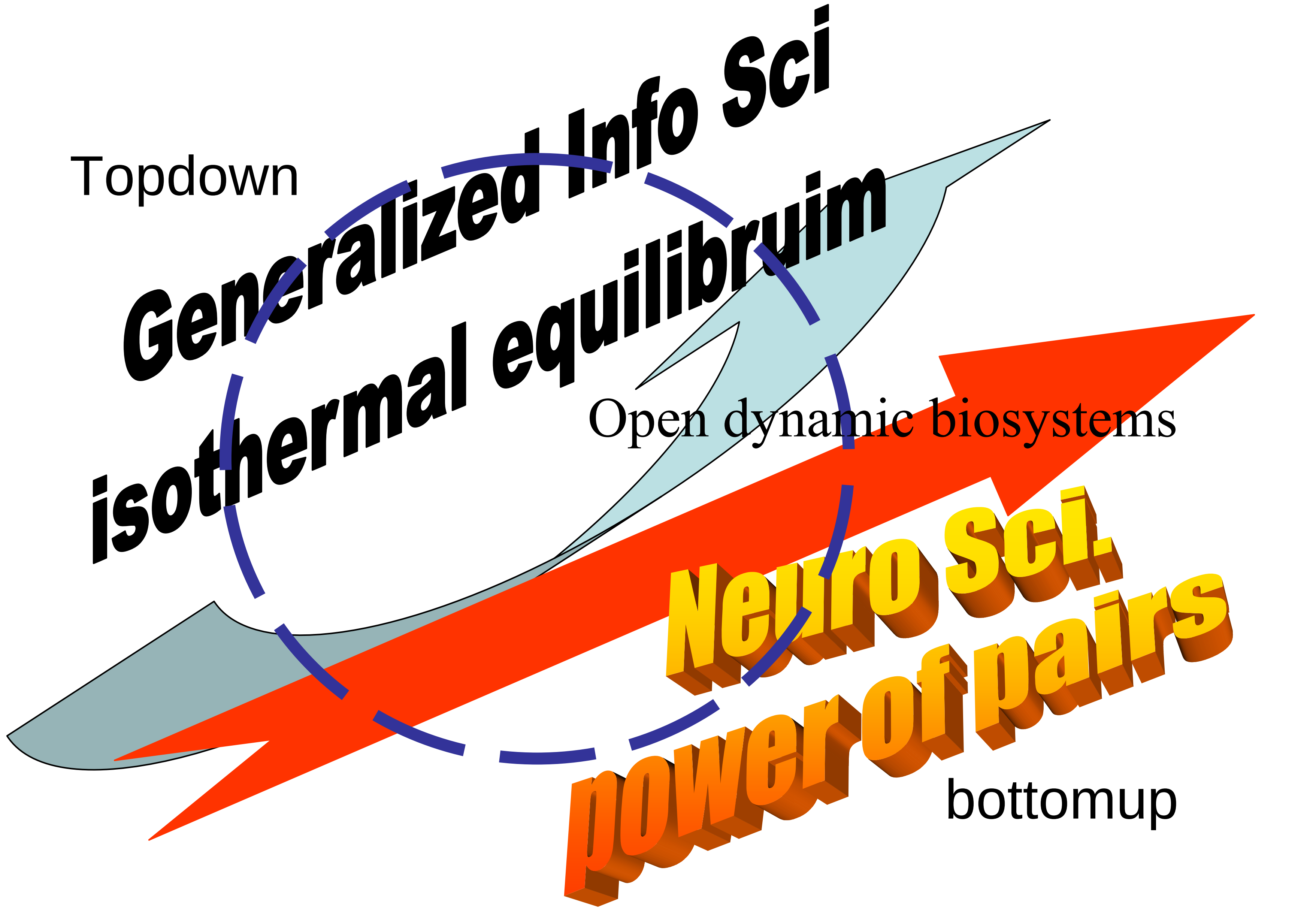
Program Officer, ONR Code 312: Sensors Surveillance and Communication ***szuh@onr.navy.mil***

1. **Book: 6th Gen Brain-style Computing, H. Szu –Peking Univ (1st introduced ANN to China, via Prof. T. Lee 1988)**
2. **Book: Single Pixel BSS, H. Szu—Fudan Univ 2004 (Shanghai S&E Publ.)**
3. **Book Chapter Unsupervised Learning ANN," H. Szu Ch. 16 CI-Experts Speaks (Ed. D. Fogel & Rabinson.) 2003.**
4. **H. Szu, et al. "Unsupervised learning with stochastic gradient"Neurocomputing, Vol 68,, pp. 130-160 Oct 2005.**
5. **H. Szu, "Thermodynamics Energy for both supervised and unsupervised learning neural nets at a constant temperature,"** ***Int'l J. Neural Sys*****. Vol.9, pp. 175-186, June 1999.**
6. **H. Szu, et al, "Landsat Spectral Demixing a la Super-resolution of Blind Matrix Inversion by Constraint MaxEnt Neural Nets," SPIE Proc. 3078, Wavelet Applications IV, pp.147-160, Orlando** ***April 1997*****.**
7. **H. Szu, et al. "NL BSS Single Pixel," 0181373 US Patent Trade Office Sept 16, 2004, pp. 700-722**

Topics

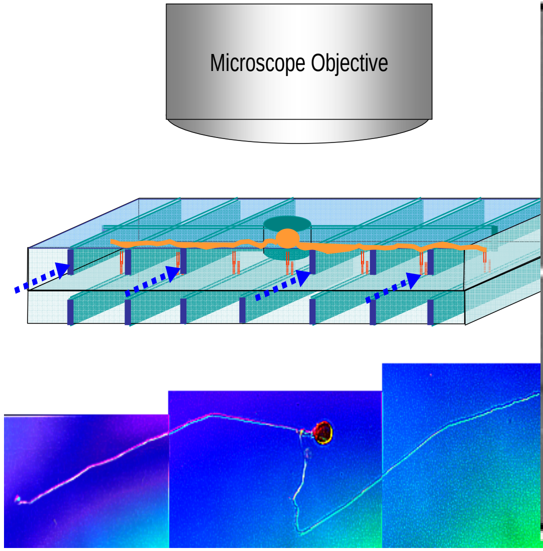
1. Generalized Brain Information Theory
   beyond Shannon Entropy **S** Concept for real world open dynamic energy exchanging ***E*** equilibrium systems for power of pairs, e.g. in legacy electrical power-line, basis of unsupervised learning Theory & Neural Annealing.
2. Real World Applications:
   Remote Sensing and Breast Tumors, mini-UAVs swarm
3. Nano-Robot Engineering: a soft-hard multiplexing the cantilever of G. Binning's ***Atomic Force Microscope (AFM)*** can build biomimetic dual color fovea by 1-D Carbon NanoTube for its QM band-gap low noise characteristics for early tumor detection by Dual Color IR Spectrum Cameras---replacing Hg thermometers
4. New R/D Brian-Central Info Sci. & Tech.

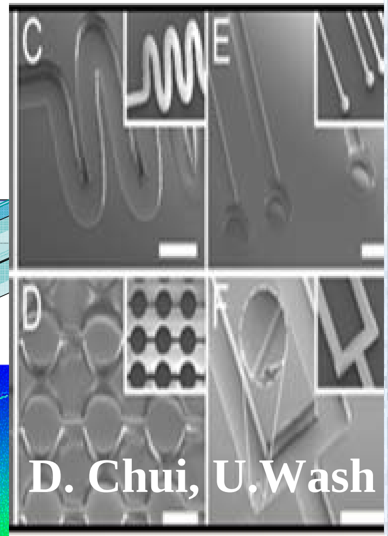
Learning is a hallmark of natural intelligence (NI), understanding of unsupervised is a key breakthrough

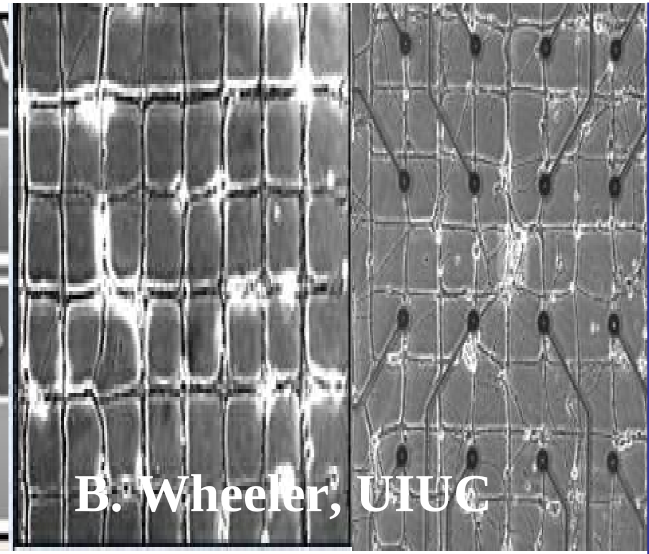


NSF Brain Science Initiative—Multidisciplinary Approach

Reverse engineering the brain:elucidating new designs, new methods of neuronal control, to tissue engineering

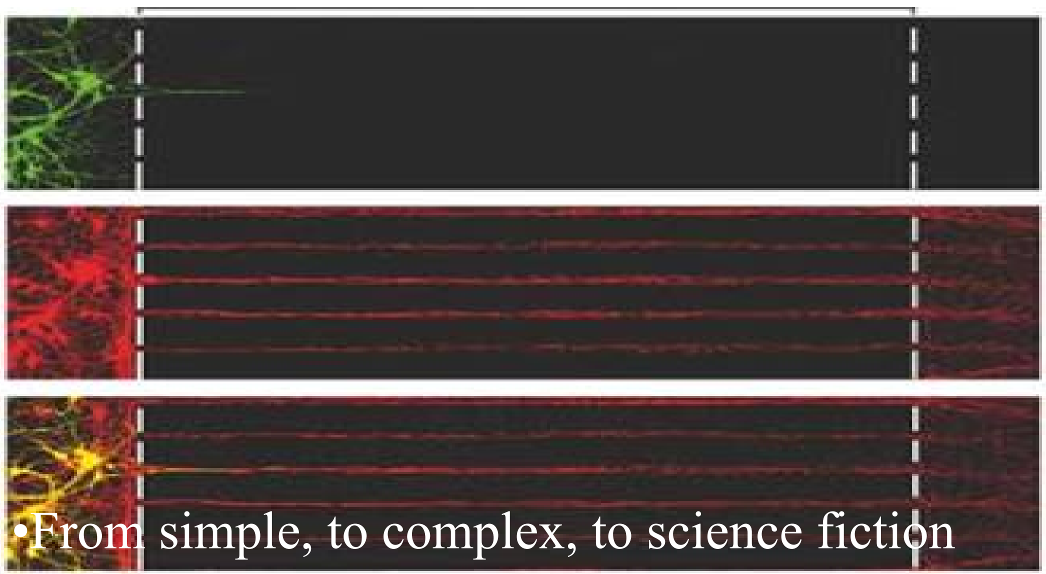




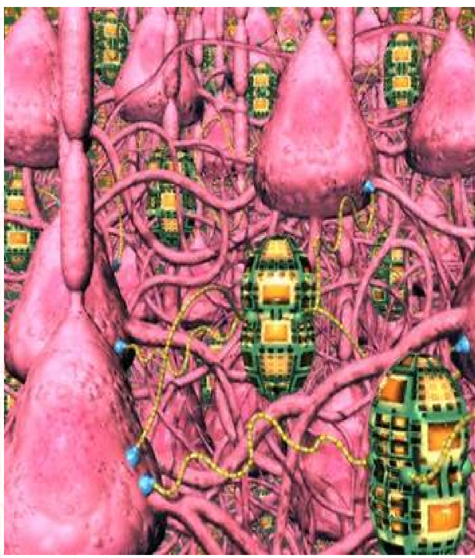


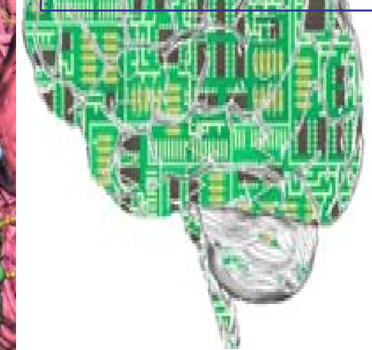
H. Szu, et. al. Live NN form 3 neurons "Peter, [..\NSF March 2007\Peter Paul Mary\VIDEO_TS\VTS_01_1.VOB](..\NSF March 2007\Peter Paul Mary\VIDEO_TS\VTS_01_1.VOB) Paul, Merry " singing together on electronic chip

Neurocomp. V6, pp.551-564, 1994



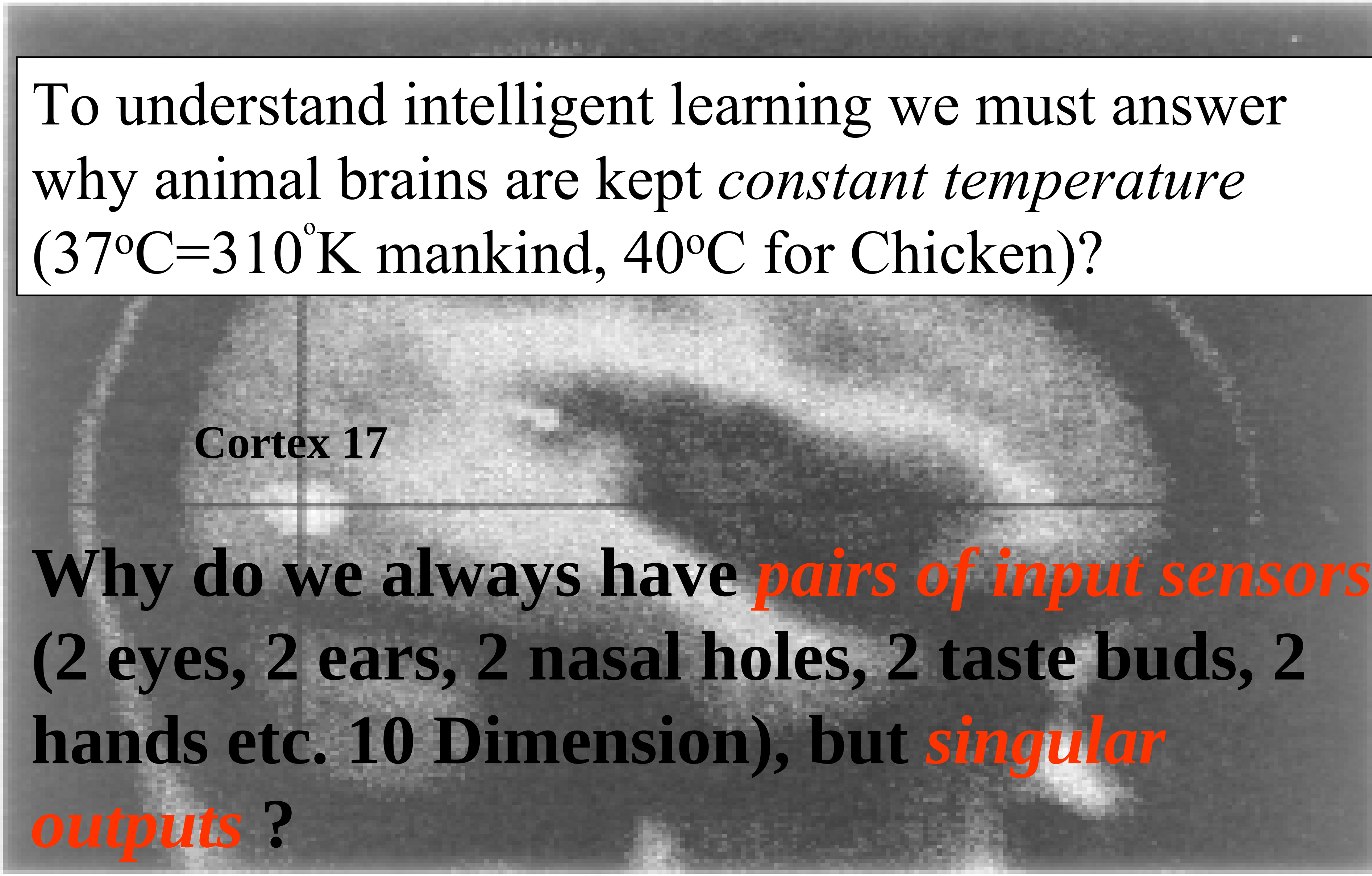
•From simple, to complex, to science fiction





www.nanotech-now.com/Art Gallery/tim-fonseca.htm

Noo Li Jeon's group, Nature Methods, **2**, 599 , (2005) To understand intelligent learning we must answer why animal brains are kept *constant temperature* ($37^{o}C = 310^{o}K$ mankind, $40^{o}C$ for Chicken)?

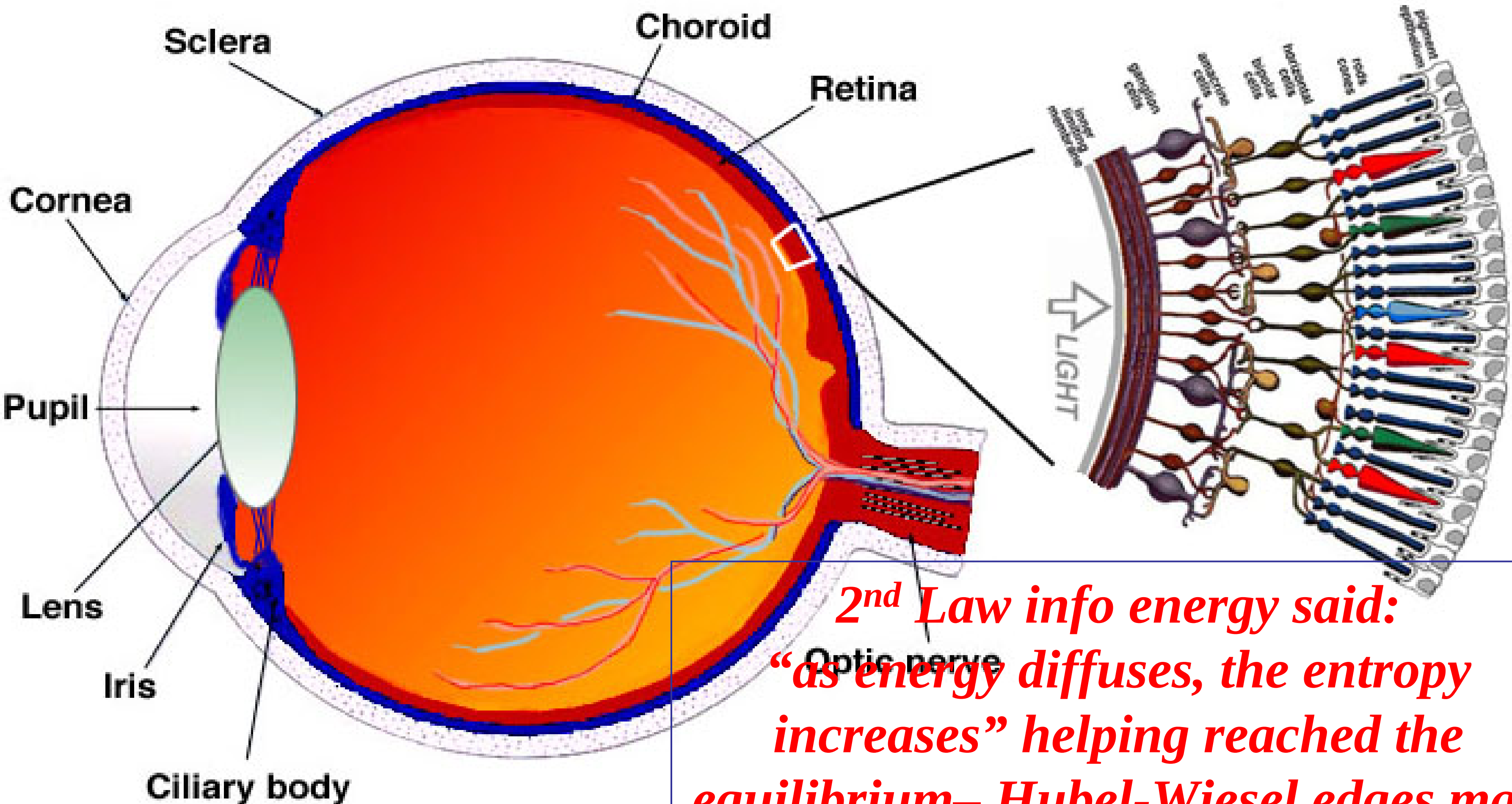


Why do we always have *pairs of input sensors* (2 eyes, 2 ears, 2 nasal holes, 2 taste buds, 2 hands etc. 10 Dimension), but *singular outputs* ?

Busy brains. Unlike low scorers on a spatial test, high scorers show a correlation between activity in a visual area (cross-hairs) and activity in the frontal lobe.

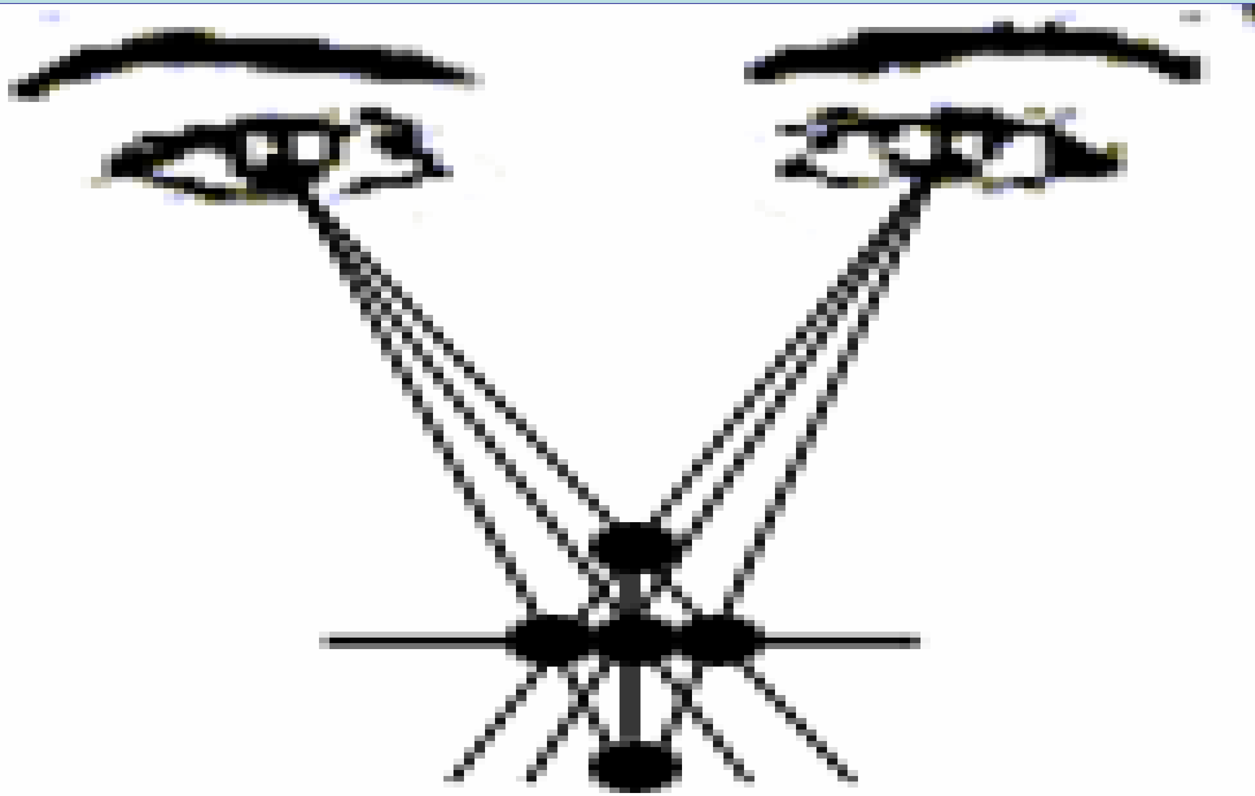
Nano-photonics must learn from our eyes?

156 millions rods-cones excitations must decay to save energy and make room for further excitations. Pseudo-1-D cone suffers less noise, enjoy single photon detection without Liquid Nitrogen coolant =Nano-photonics need CAD NanoRobot manipulator, from 1 meter to 10^{-9} meter!



2^{nd} Law info energy said: "as energy diffuses, the entropy increases" helping reached the equilibrium– Hubel-Wiesel edges map

Why do we have two eyes?
Why receptors input prefer pairs, while emitters are singles?



Answers:

(1)Hardware fault tolerance;

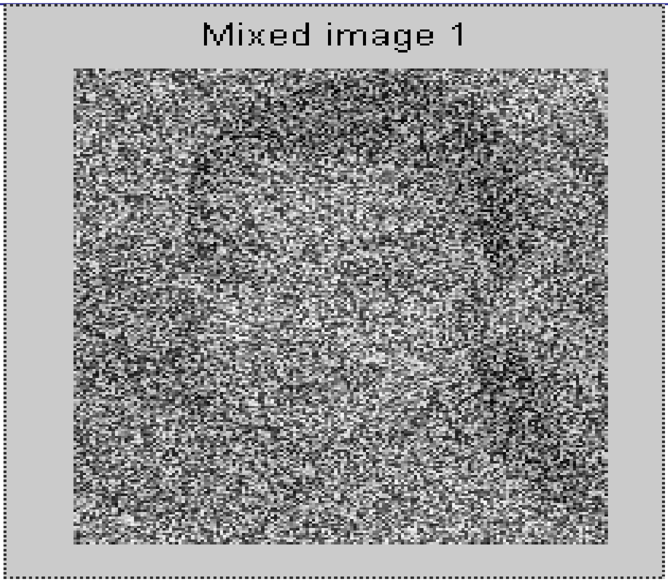
(2)Stereovision—no range info David Marr Binocular Paradox

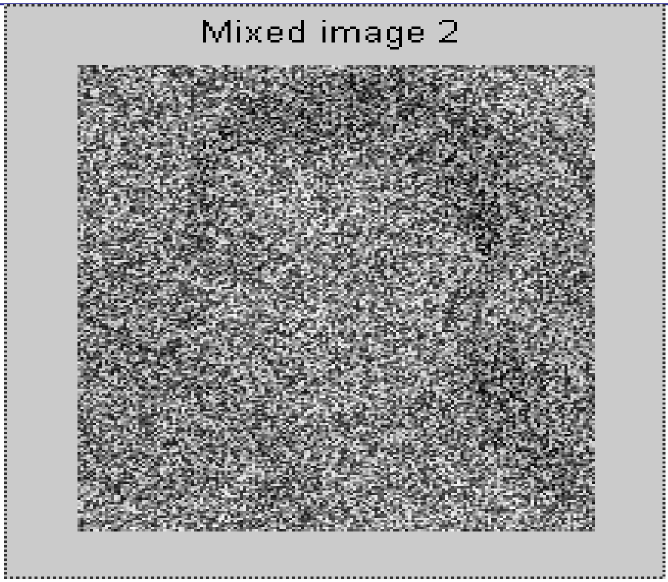
(3)Unsupervised learning by Coincidence Account

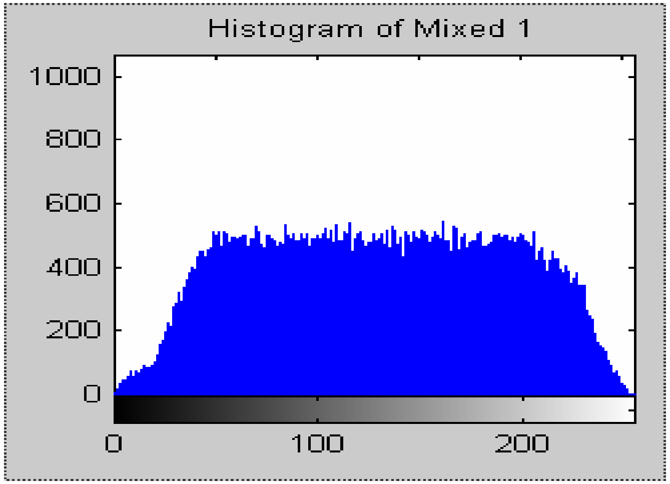
(4) All above (correct answer)

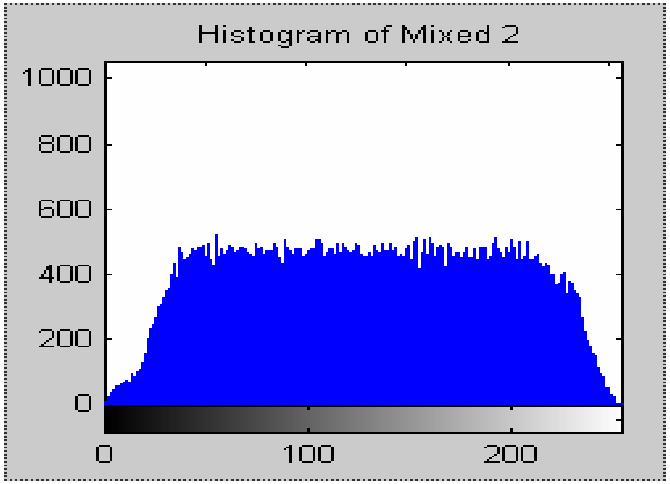
Two eyes could see through the fog---

two-eye a perfect restoration as opposed to one eye image processing is merely re-shoveling snow!

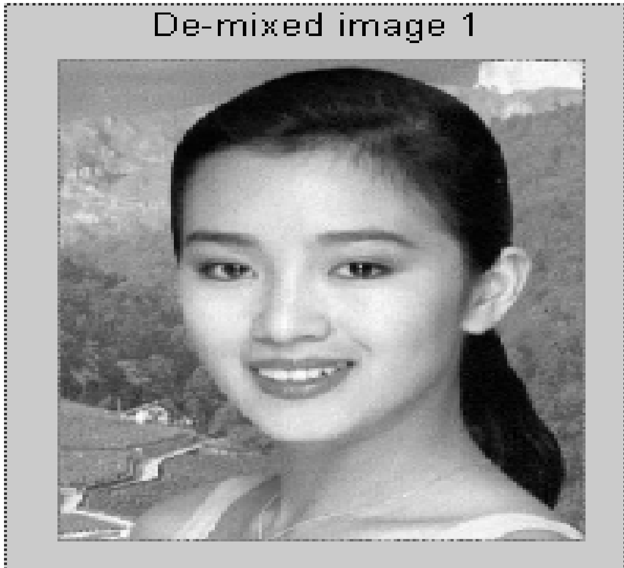


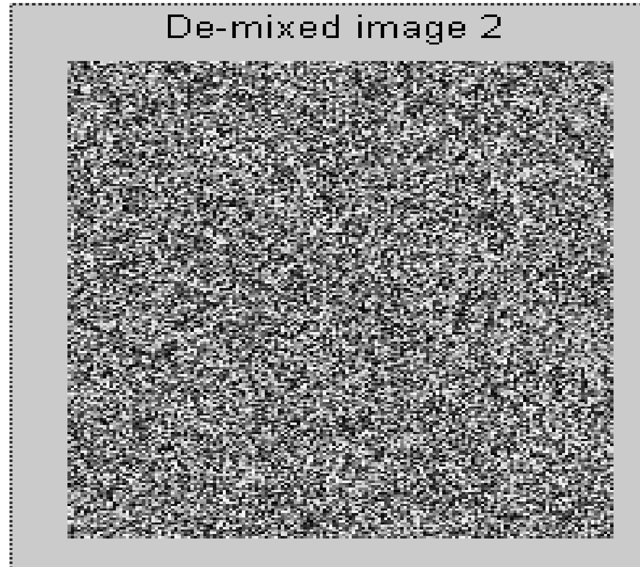


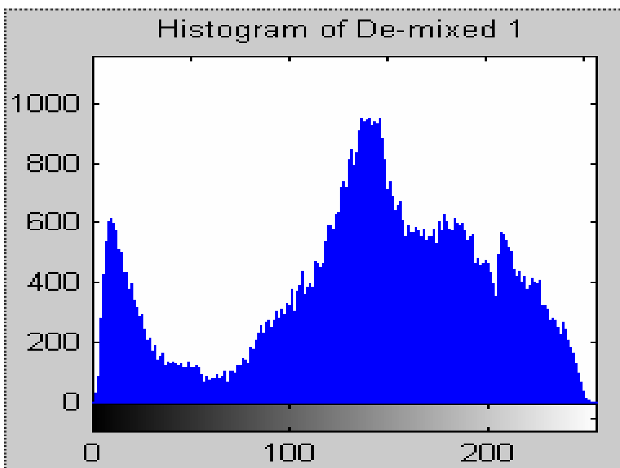




While agreements must be signals, disagreements, noises, that are universal and need no teacher!!







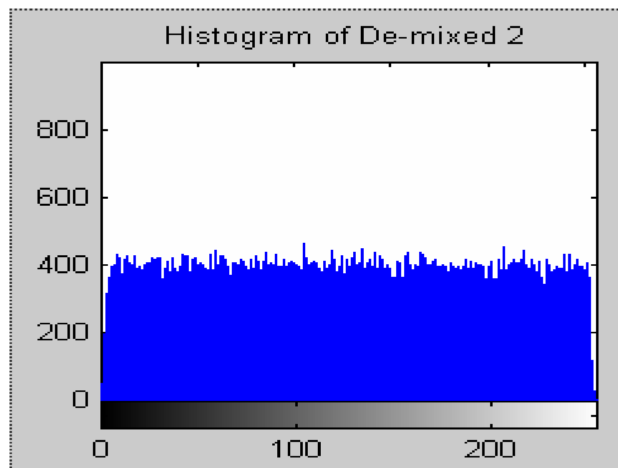
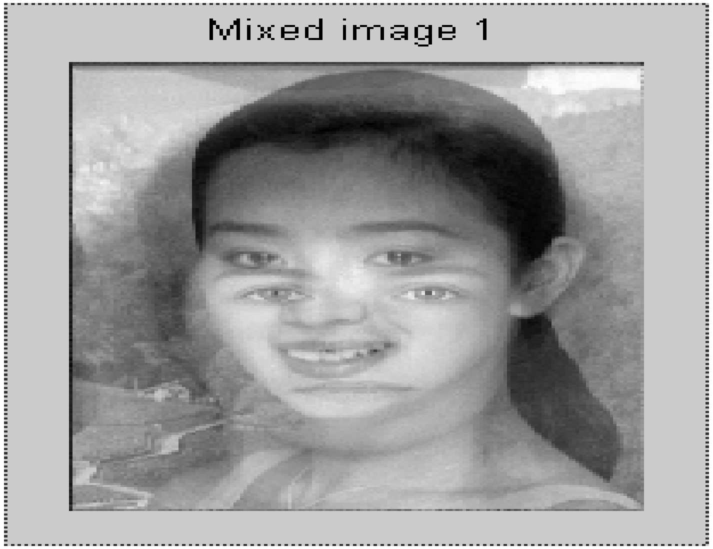
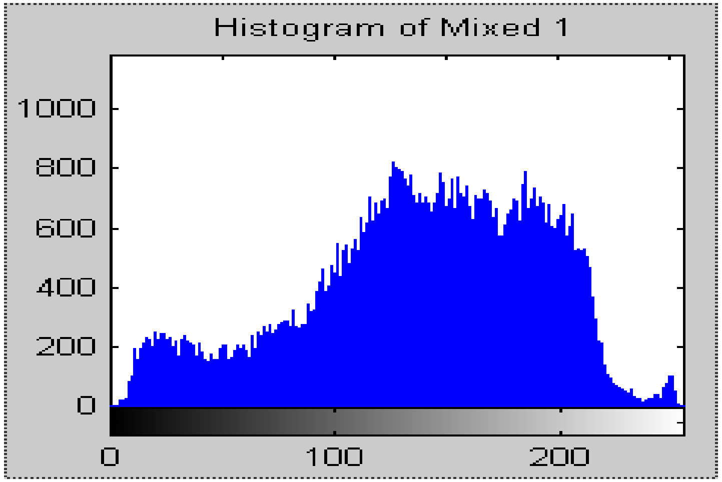
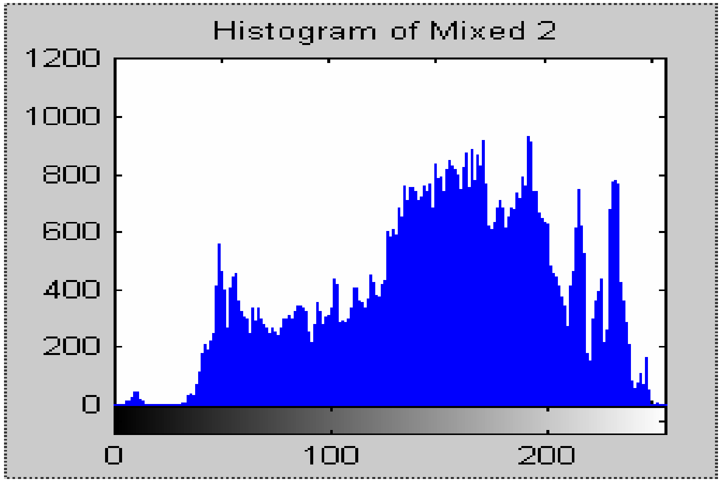


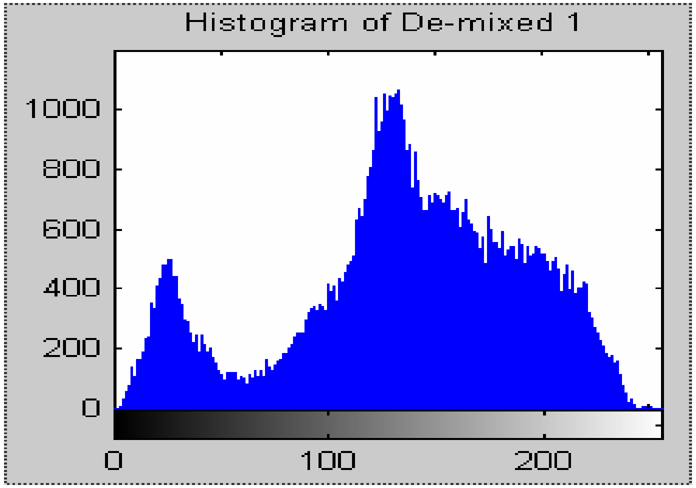
Illustration of Blind Source Separation

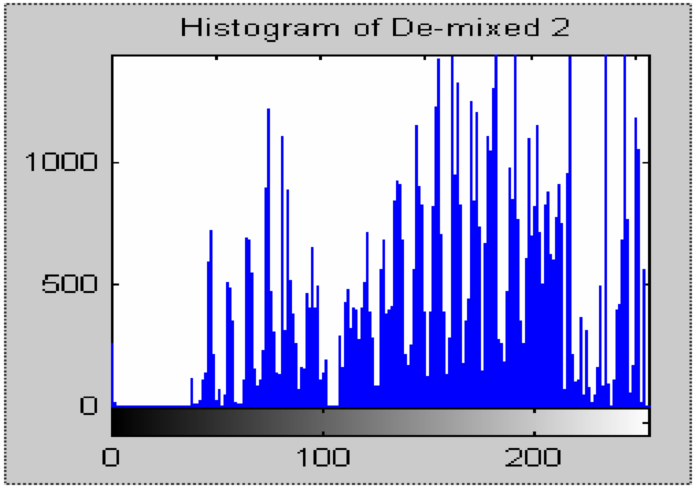






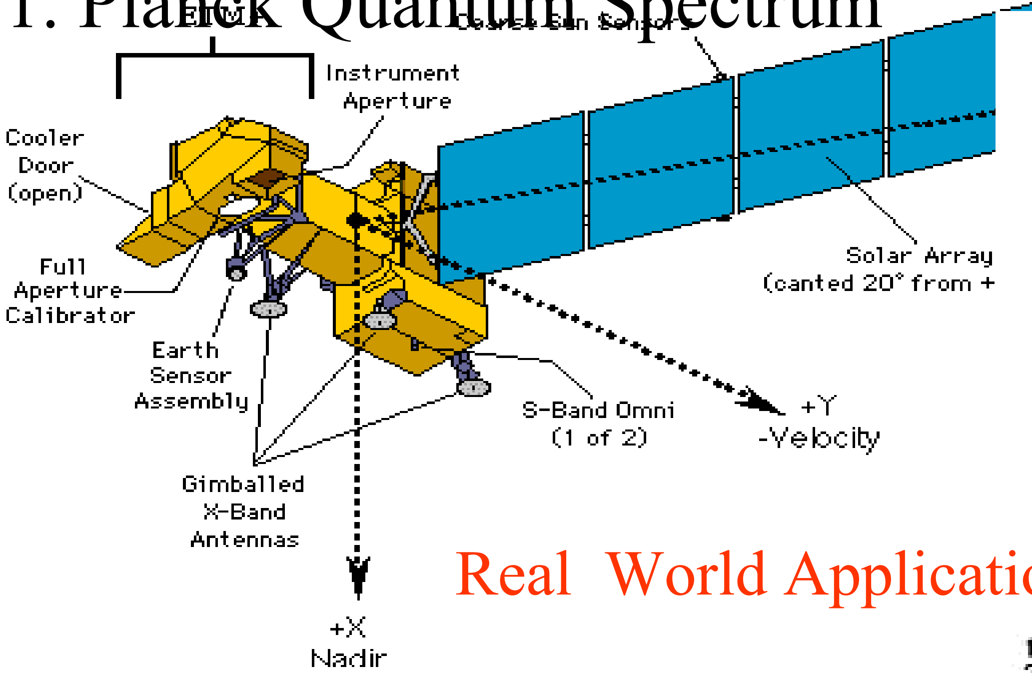
De-mixed images and histograms

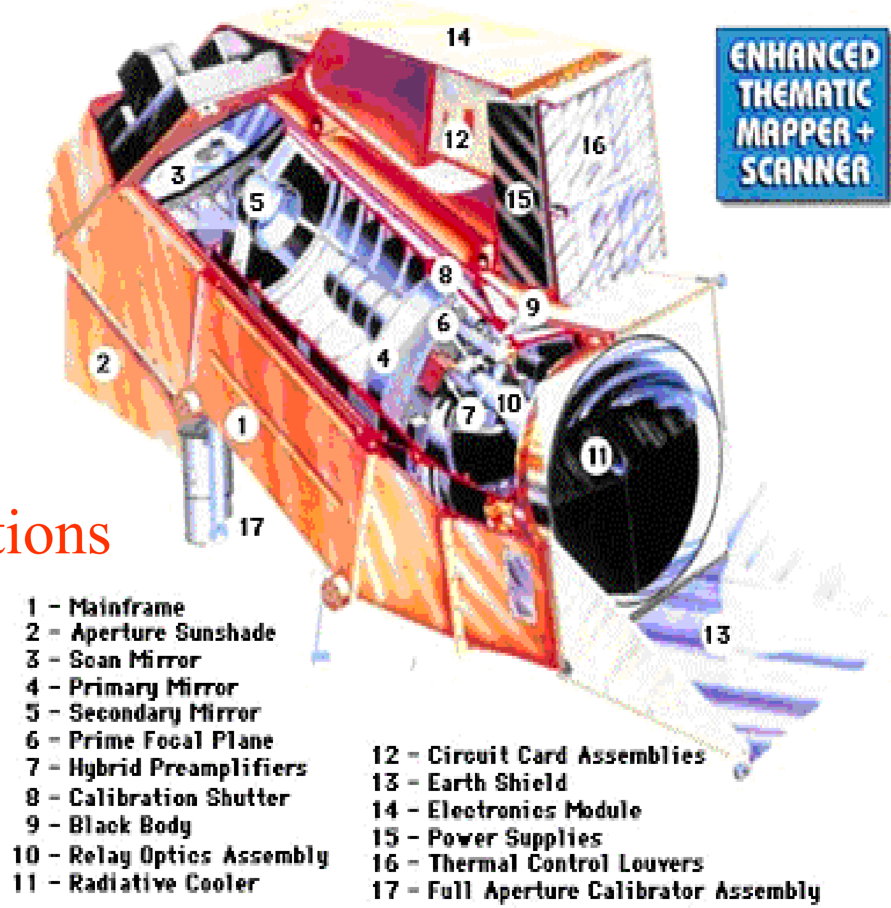




Translation from military surveillance to home care
From tank to Tumor ("***NL Blind Demixing Single Pixel etc***," ***Szu et al. US Patent 0181375, 2004, Sept 15***)

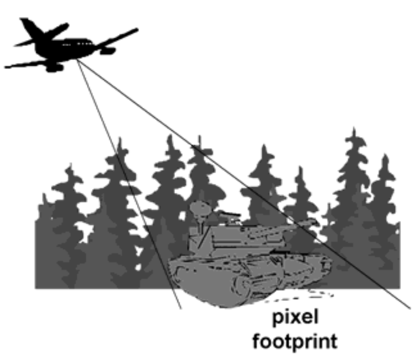
1. Planck Quantum Spectrum

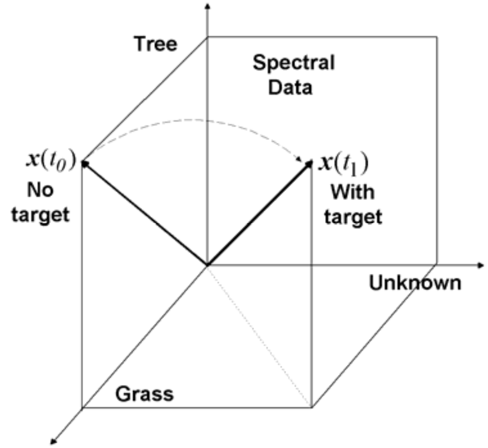


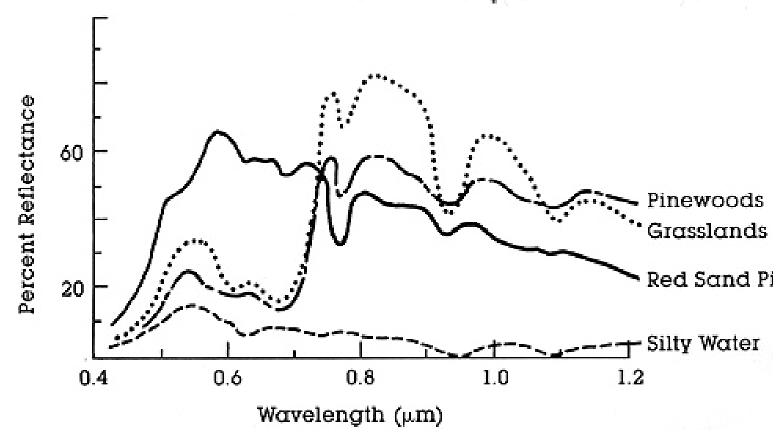


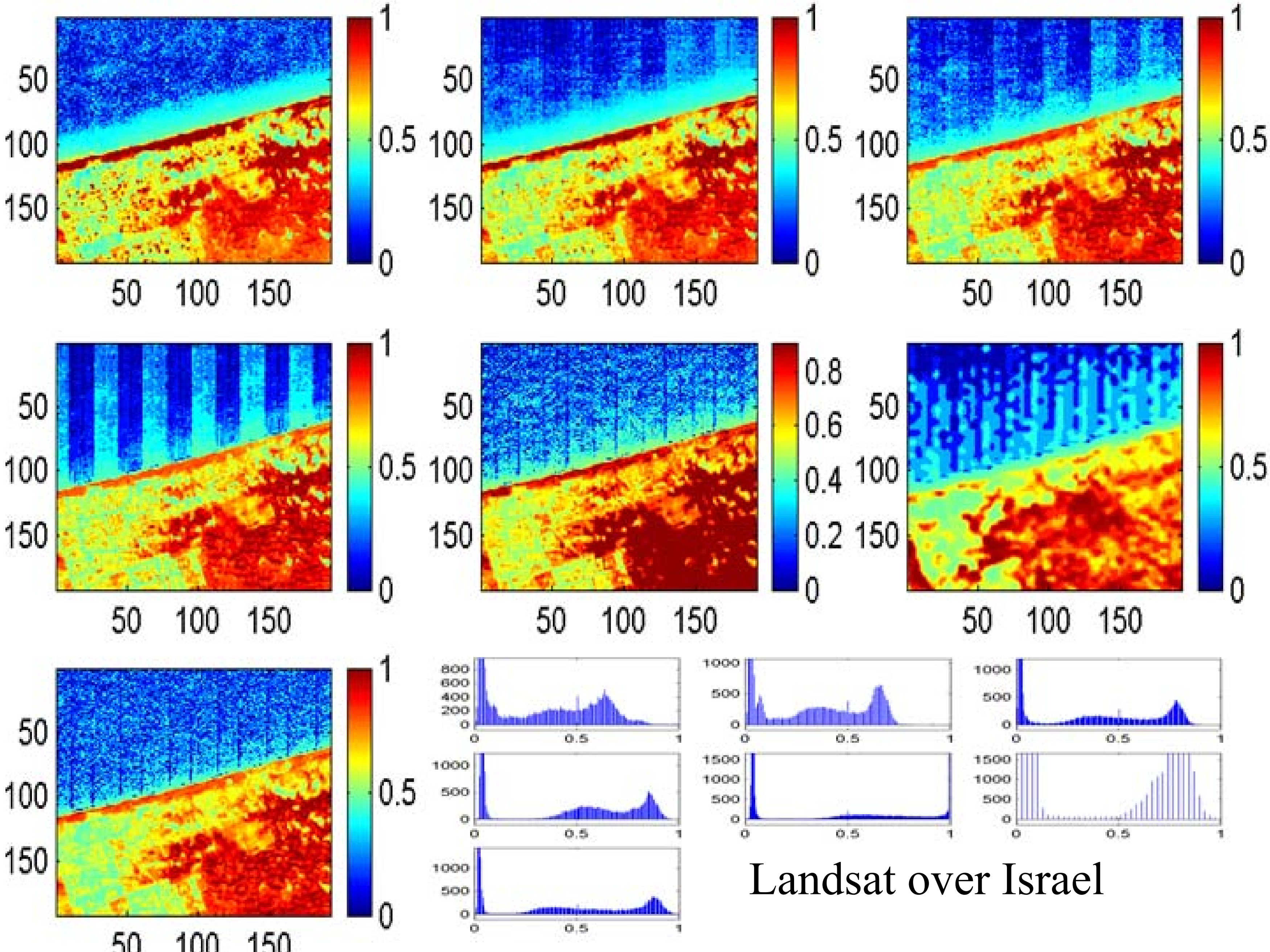
Real World Applications

TM AND ETM+ SPECTRAL BANDWIDTHS								
Bandwidth (μ) Full Width - Half Maximum								
Sensor	Band 1	Band 2	Band 3	Band 4	Band 5	Band 6	Band 7	Band 8
TM	0.45 - 0.52	0.52 - 0.60	0.63 - 0.69	0.76 - 0.90	1.55 - 1.75	10.4 - 12.5	2.08 - 2.35	N/A
ETM+	0.45 - 0.52	0.53 - 0.61	0.63 - 0.69	0.78 - 0.90	1.55 - 1.75	10.4 - 12.5	2.09 - 2.35	.52 - .90

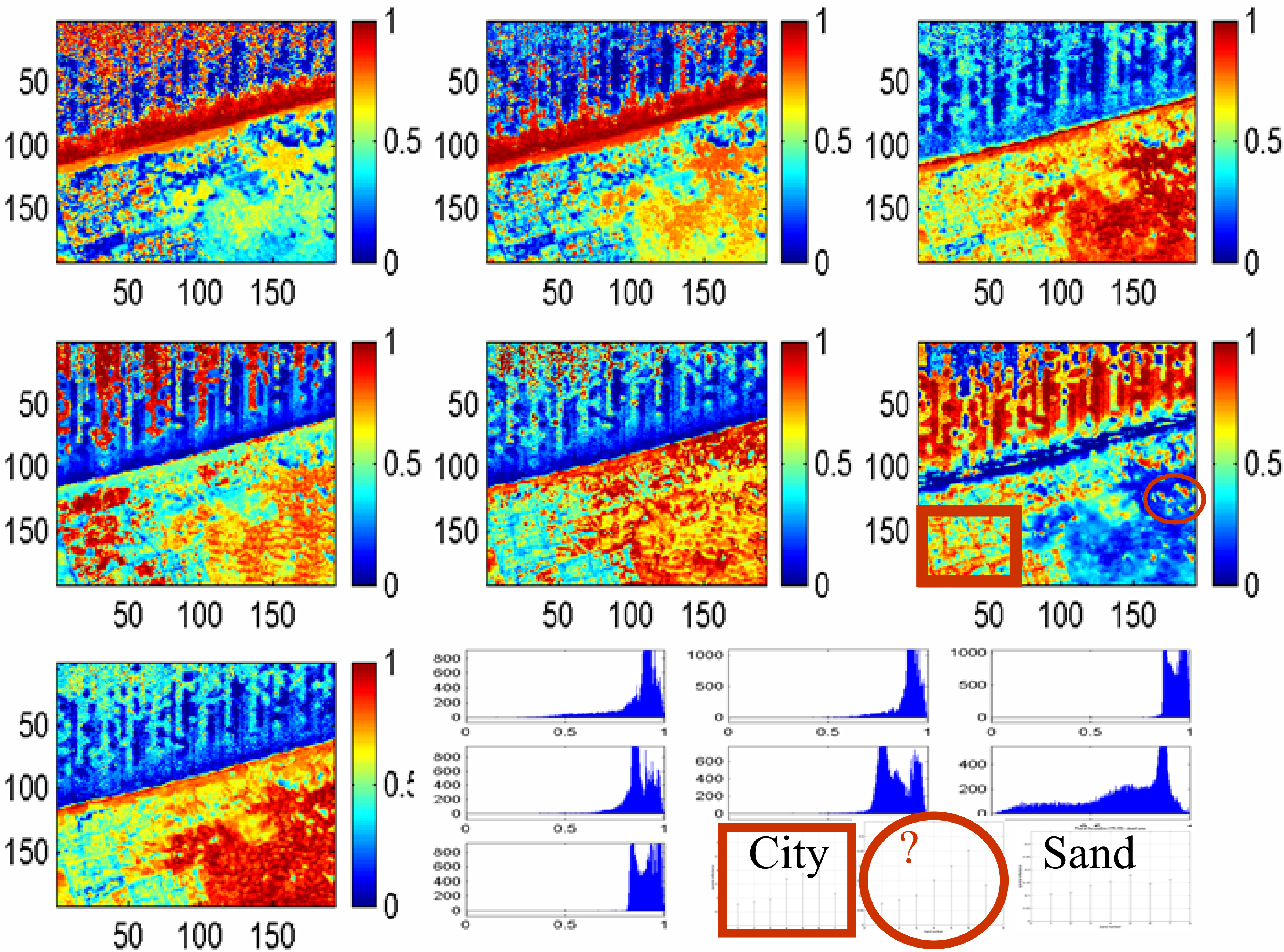








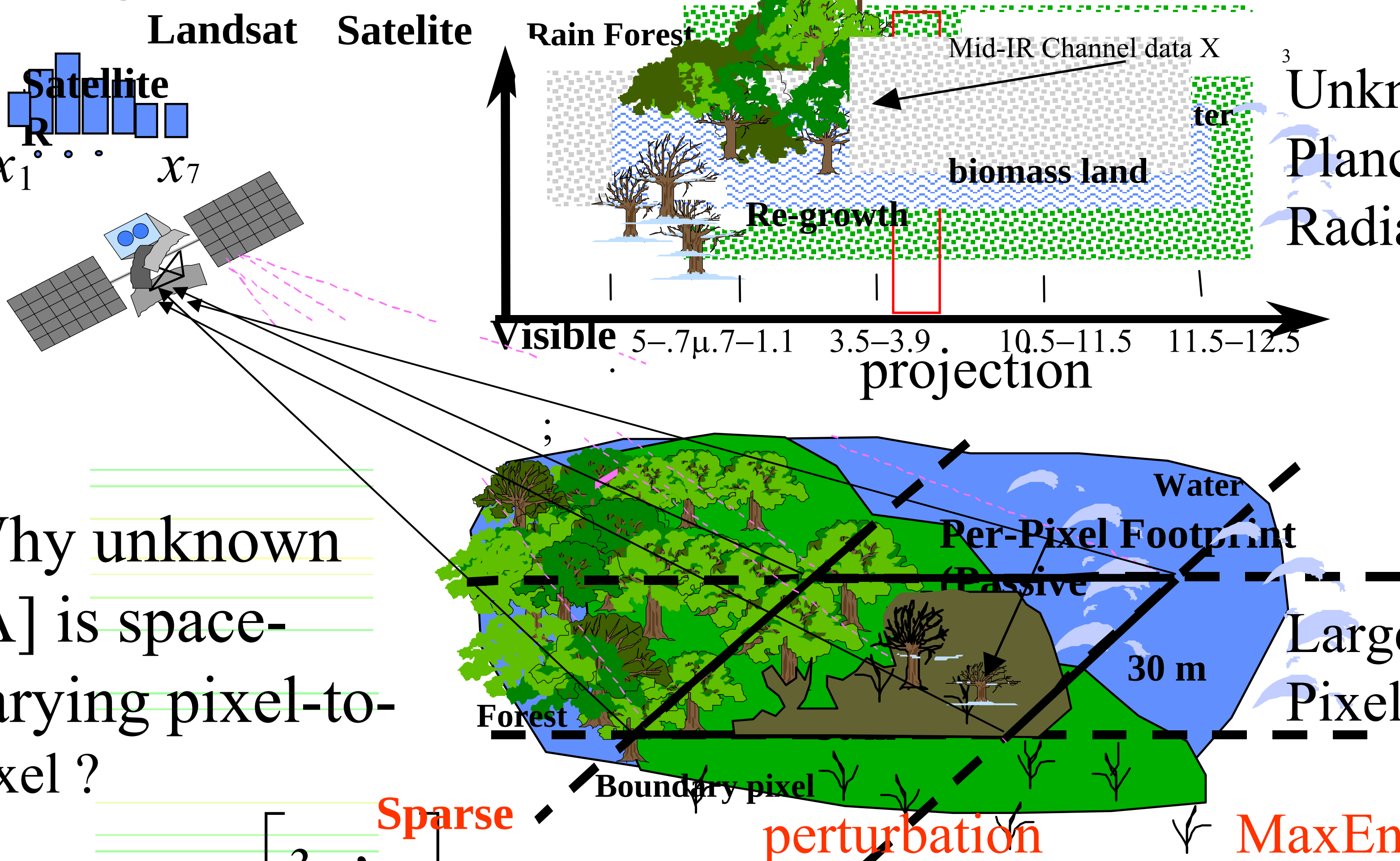
Landsat over Israel



City
?
Sand

(1) Diurnal & seasonal variations yield unknown objects-spectral table matrix [A]

(2) Large Footprint per pixel requires unknown composition Labels s per pixel



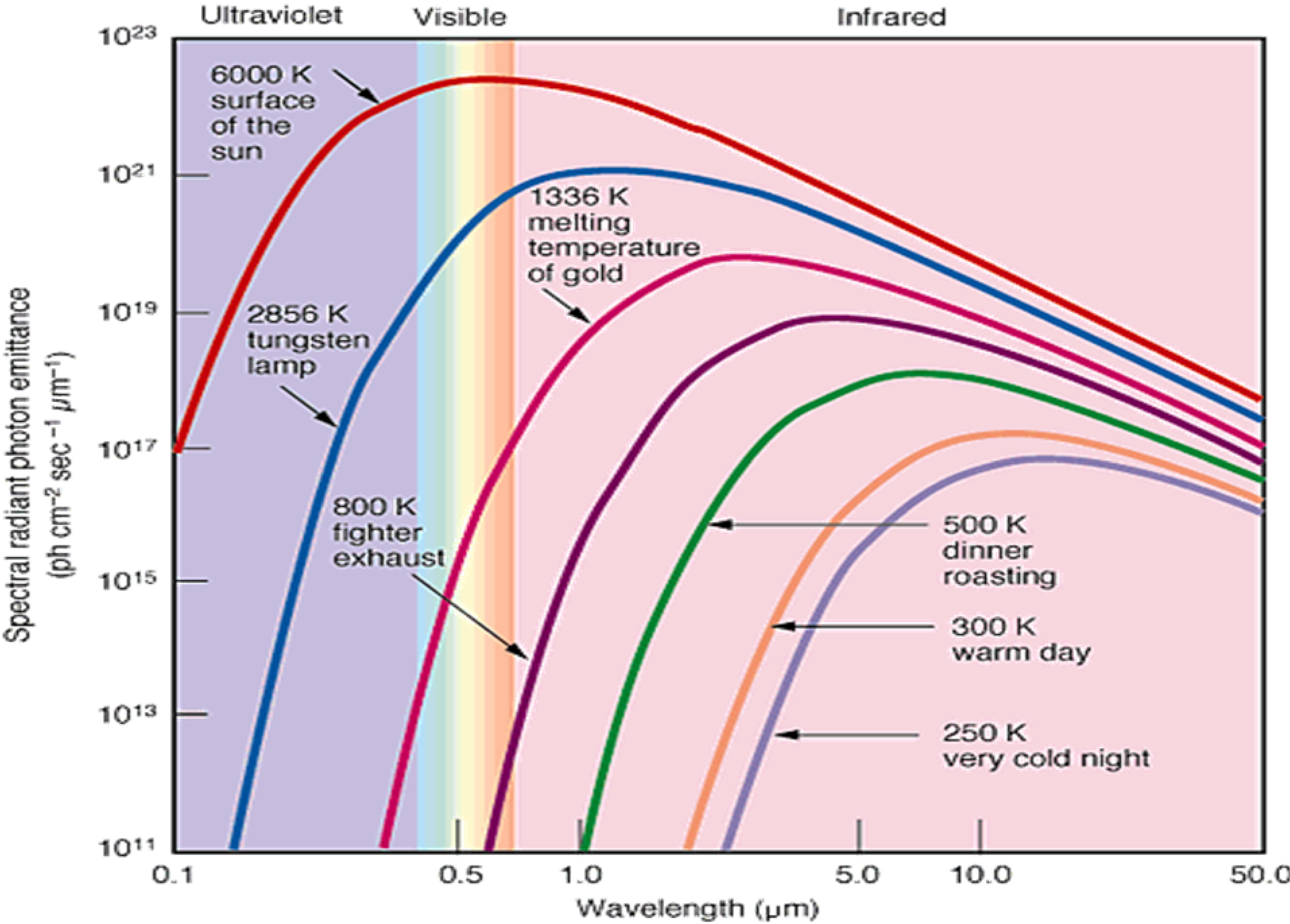
Unknown Planck Law Radiation [A]

Large Single Pixel Footprint

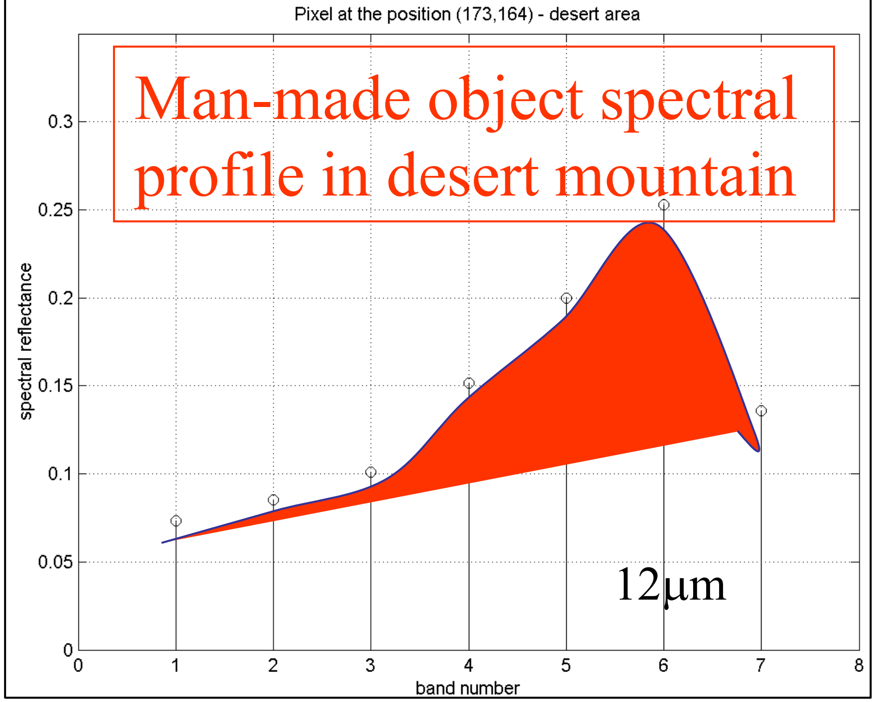
Why unknown [A] is space-varying pixel-to-pixel ?

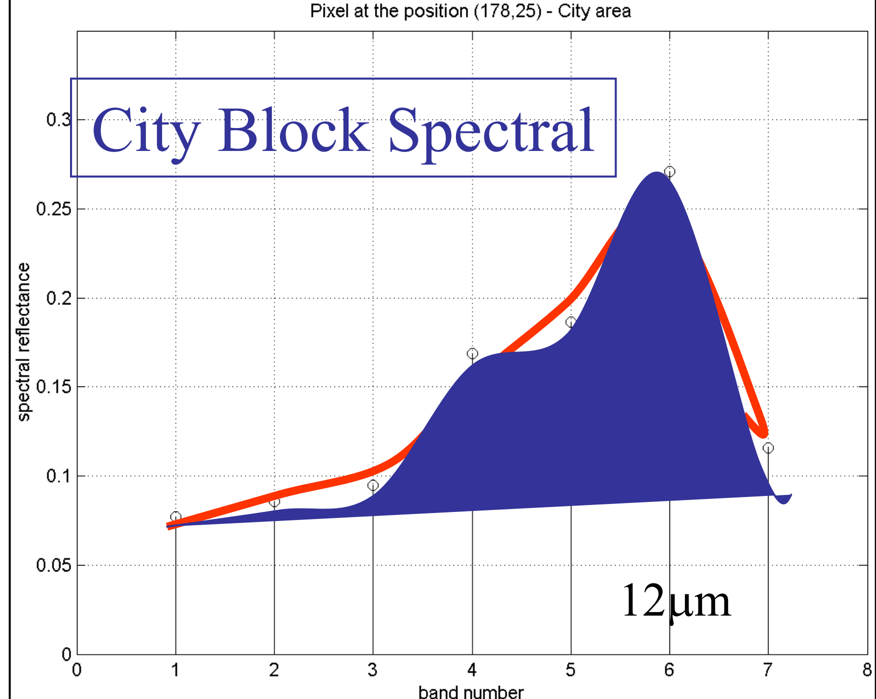
Sparse perturbation MaxEnt

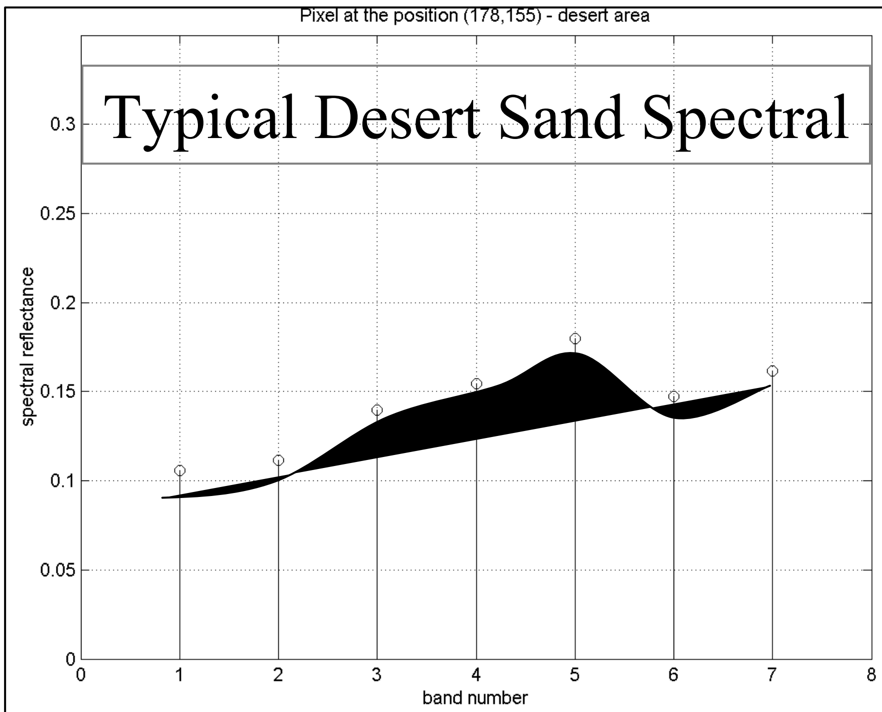
$$\mathbf{x}=\begin{bmatrix}3\\2\\1\end{bmatrix}=\begin{bmatrix}3 & a' & a\\2 & a' & a\\1 & a' & a\end{bmatrix}\begin{bmatrix}1\\0\\0\end{bmatrix}\cong\begin{bmatrix}3+\delta' & 3-\delta' & a\\2+\delta'' & 2-\delta'' & a\\1+\delta''' & 1-\delta''' & a\end{bmatrix}\begin{bmatrix}1/2+\varepsilon\\1/2-\varepsilon\\0\end{bmatrix}=\begin{bmatrix}3 & 3 & 3\\2 & 2 & 2\\1 & 1 & 1\end{bmatrix}\begin{bmatrix}1/3\\1/3\\1/3\end{bmatrix}$$



Ultraviolet
Visible
Infrared
10^{23}
10^{21}
10^{19}
10^{17}
10^{15}
10^{13}
10^{11}
Spectral radiant photon emittance
(ph cm^{-2} sec^{-1} μm^{-1})
6000 K surface of the sun
2856 K tungsten lamp
1336 K melting temperature of gold
800 K fighter exhaust
500 K dinner roasting
300 K warm day
250 K very cold night
0.1
0.5
1.0
5.0
10.0
50.0
Wavelength (μm)



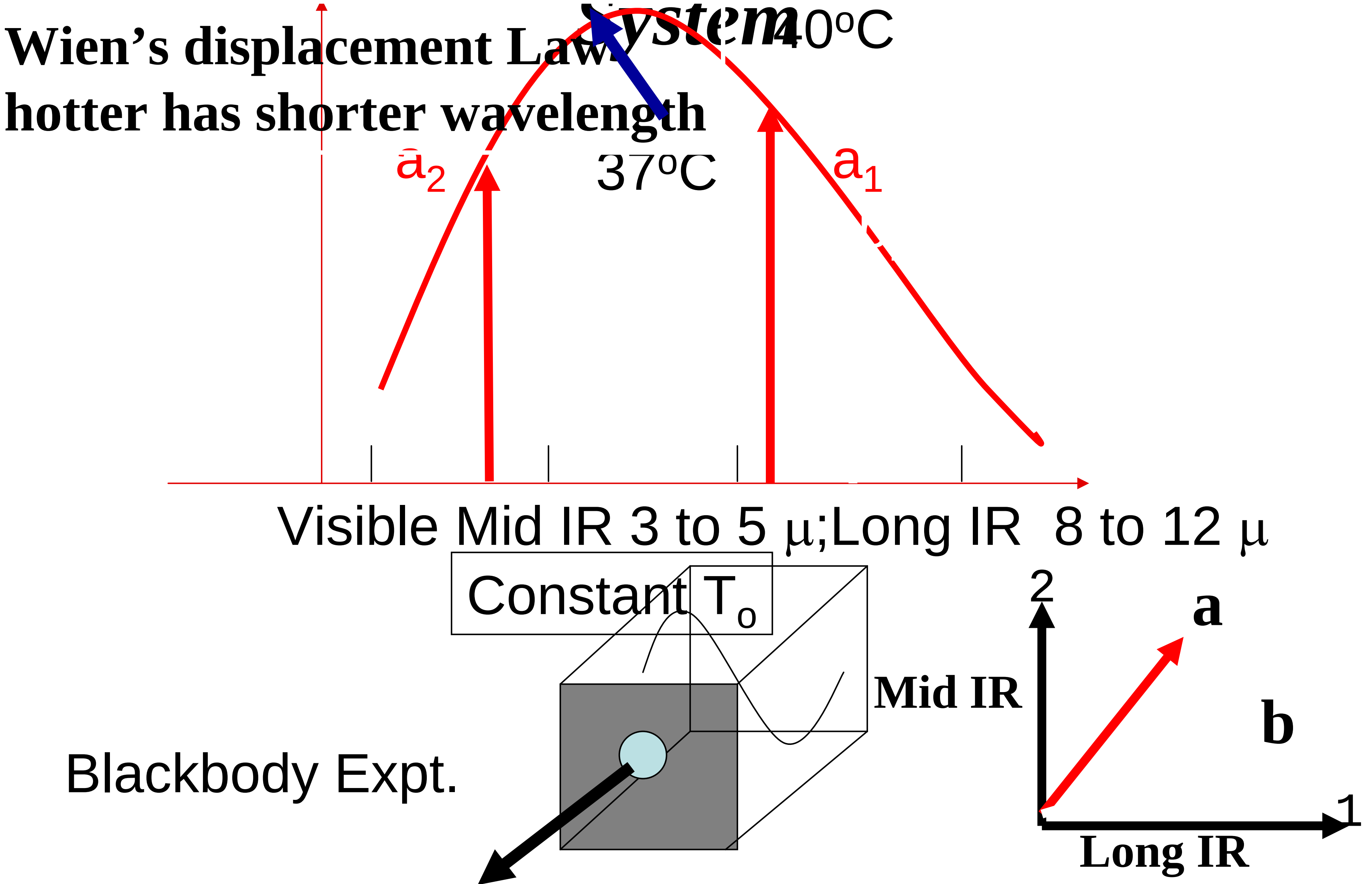




In Duda-Hart sense of unlabelled Data X LCNN discovers Planck Radiation Law as overlapping Fuzzy Memberships

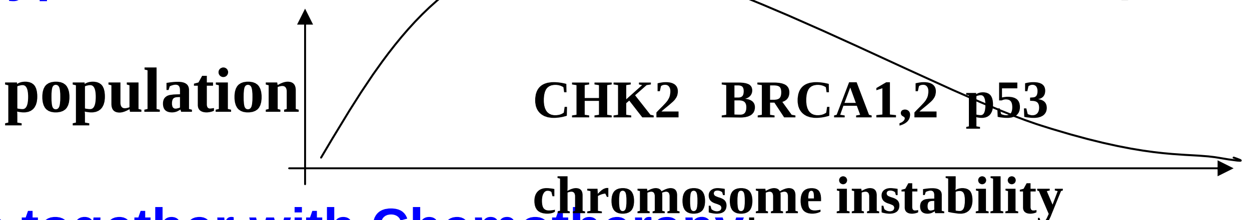
Top left pixel (173,164): LCNN discovered a "man-made construction" in desert mountain area because of the correspondence pixel (178, 25) in City block shown Top right; whereas a "typical" pixel (178,155) in desert shown in Bottom left.

Isothermal Equilibrium Black Body Radiation Planck Quantum Mechanics System



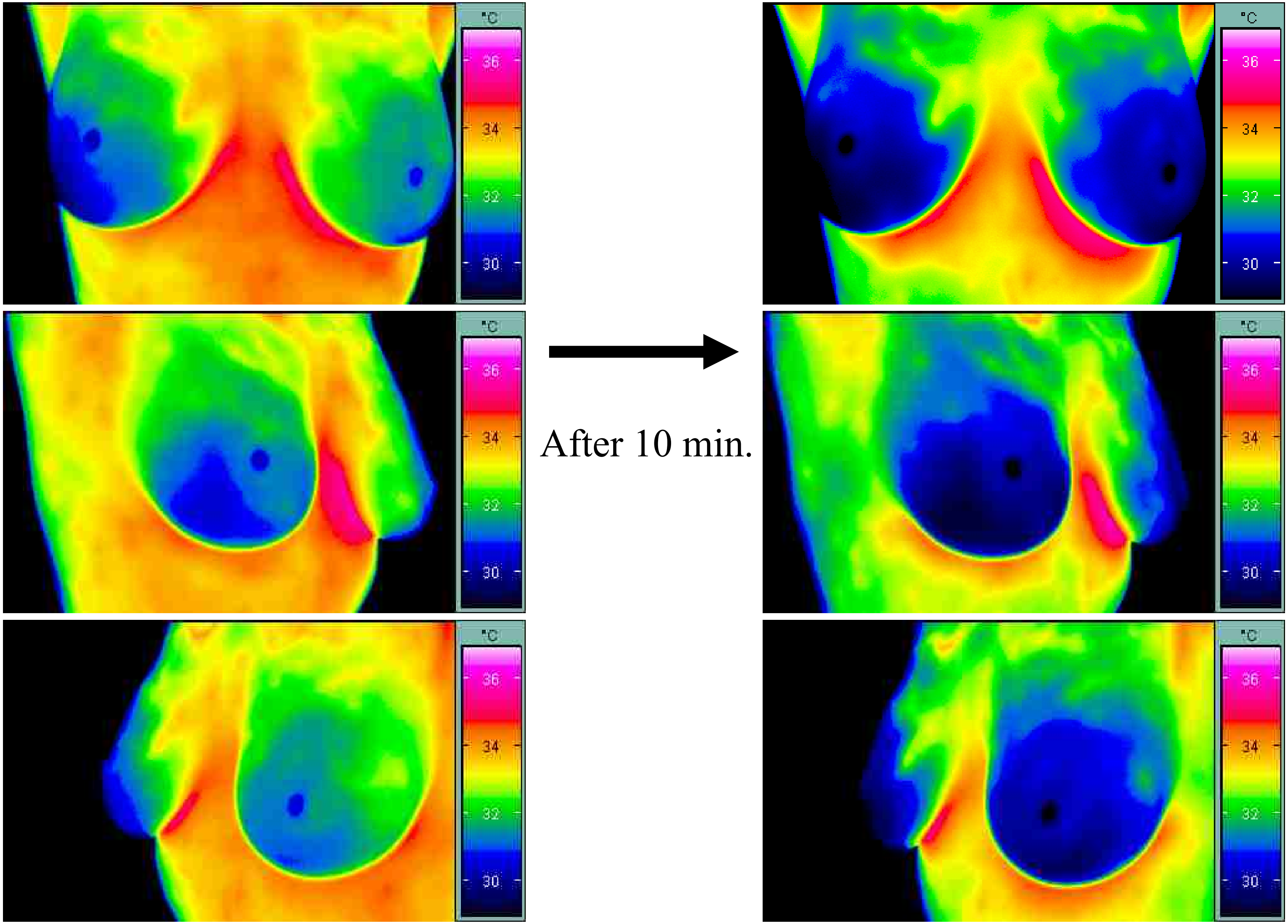
Current Status of Breast Cancers Sci. 308, p.1248

- **Phenotype**: e.g. **UV-ATR, IR-ATM**, among others, Without **protein Bim** guards selective ***apoptosis***: **initiation** [primary tumor, **invasion** intra-vasation (***angiogenesis early detectable by IR spectrogram***), extra-vasation **transport**] **metastasis**---fractal distribution, **micro-calcification (X-ray *mammogram*)**
- **Genotype:** e.g. **UV-CHK1; IR-CHK2; other BRC1, BRC2, p53**

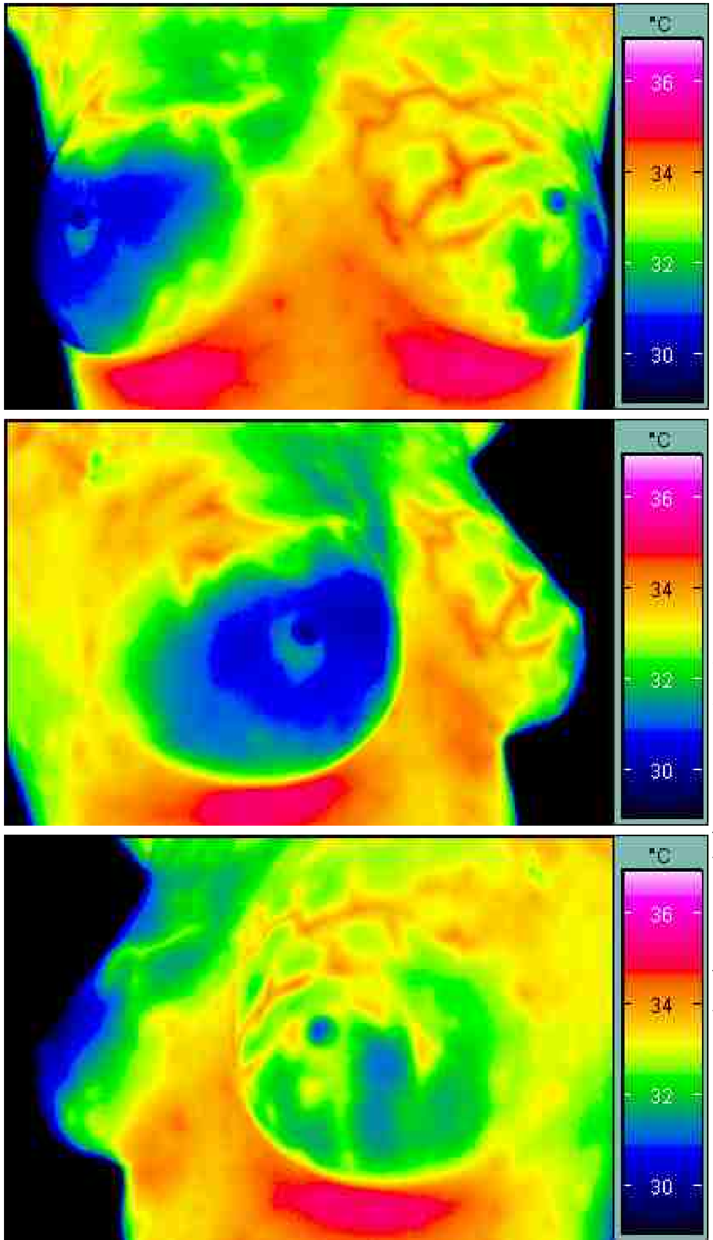


- **Drugs together with Chemotherapy**:
- **Avastin by Genetch an antibody protein injected for metastasis extended life 6 to 11 months also for colon, lung, etc.**
- **Herceptin for HER2 protein prevent 52% recurrence.**
- **Femara for hormone reduce 19% recurrence vs. tanoxifan.**
- **2nd Gen reverse angiogenesis drug targeted molecular signals path & sites**
- **Sorafenib by Bayers (**2001) binds angiogenesis promotion protein (Vascular Endothelial GF) **for kidney cancer**
- **Sutent by Pfizer (**2000 Cherrington SUGEN) **pills blocks receptor of Tyrosin kinases for adding phosphate to protein & enzymes for GIST, (Breast, Indian U., Miller) passed FDA March 2006**
- **Epidemiology**: **More beasts' than lungs' cancers**.(***220K vs 80K***); **Less mortality (19% vs. 39%). Nulliparous had high risk; Singapore postwar boosted DCIS from 0.1% to 1% attributed to phenotype lifestyle & offspring form 5 to 1 child.**

“From Tanks to Tumors” State of the Art: Healthy Breasts (shown left) and After 10 min. waiting (right) by one camera passive Thermal Scan

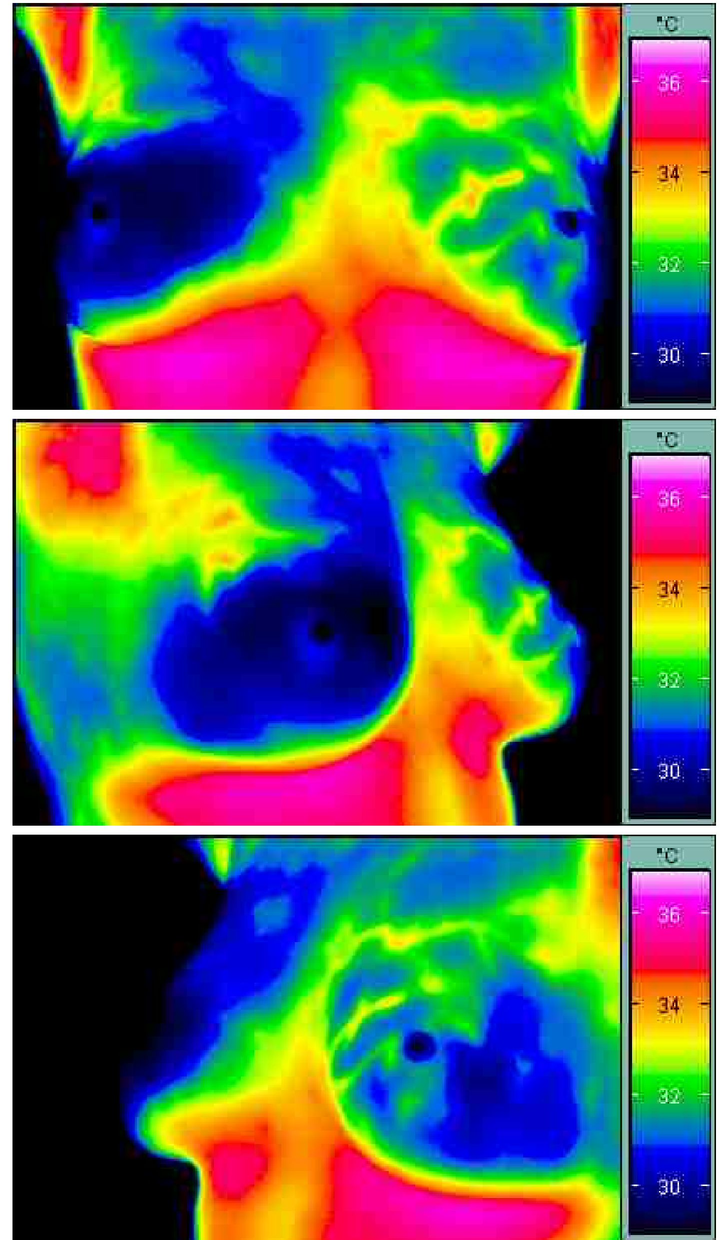


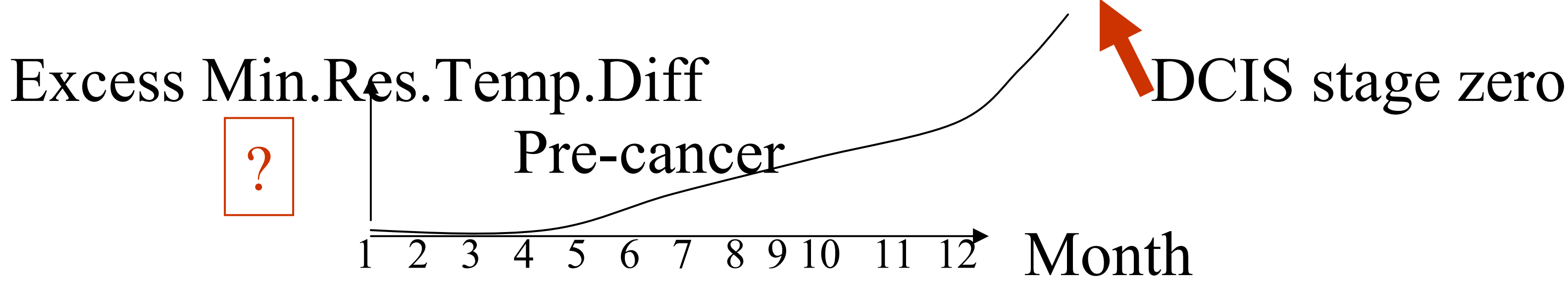
State of the Art: Pathological Breasts Before (shown left) and After 10 min. waiting (right) by one camera passive Thermal Scan (IRI)

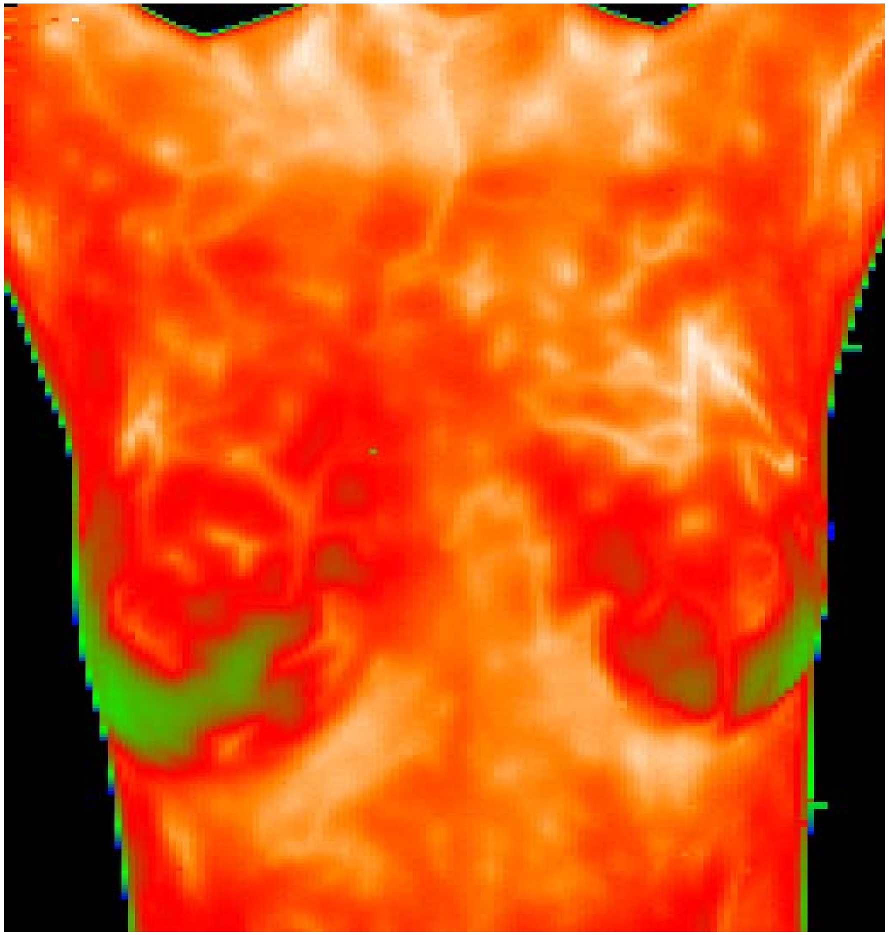


After 10 min.

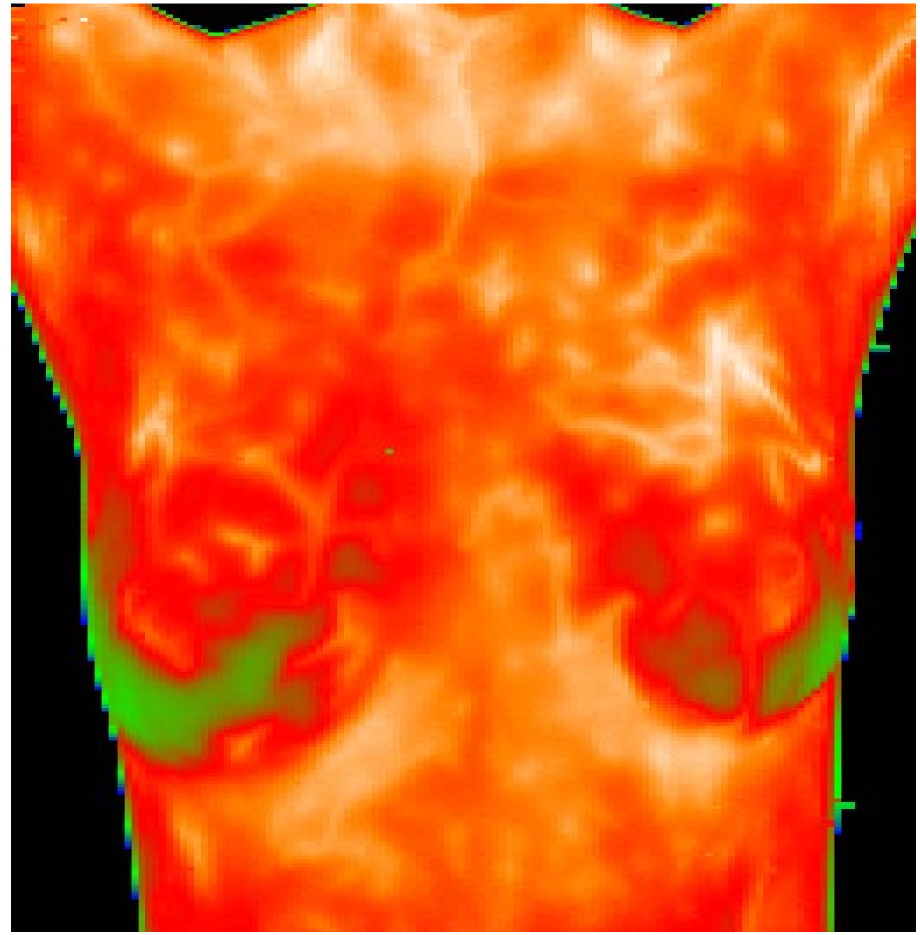
DCIS
zero stage
requires
excess
heat supply







Mid IR

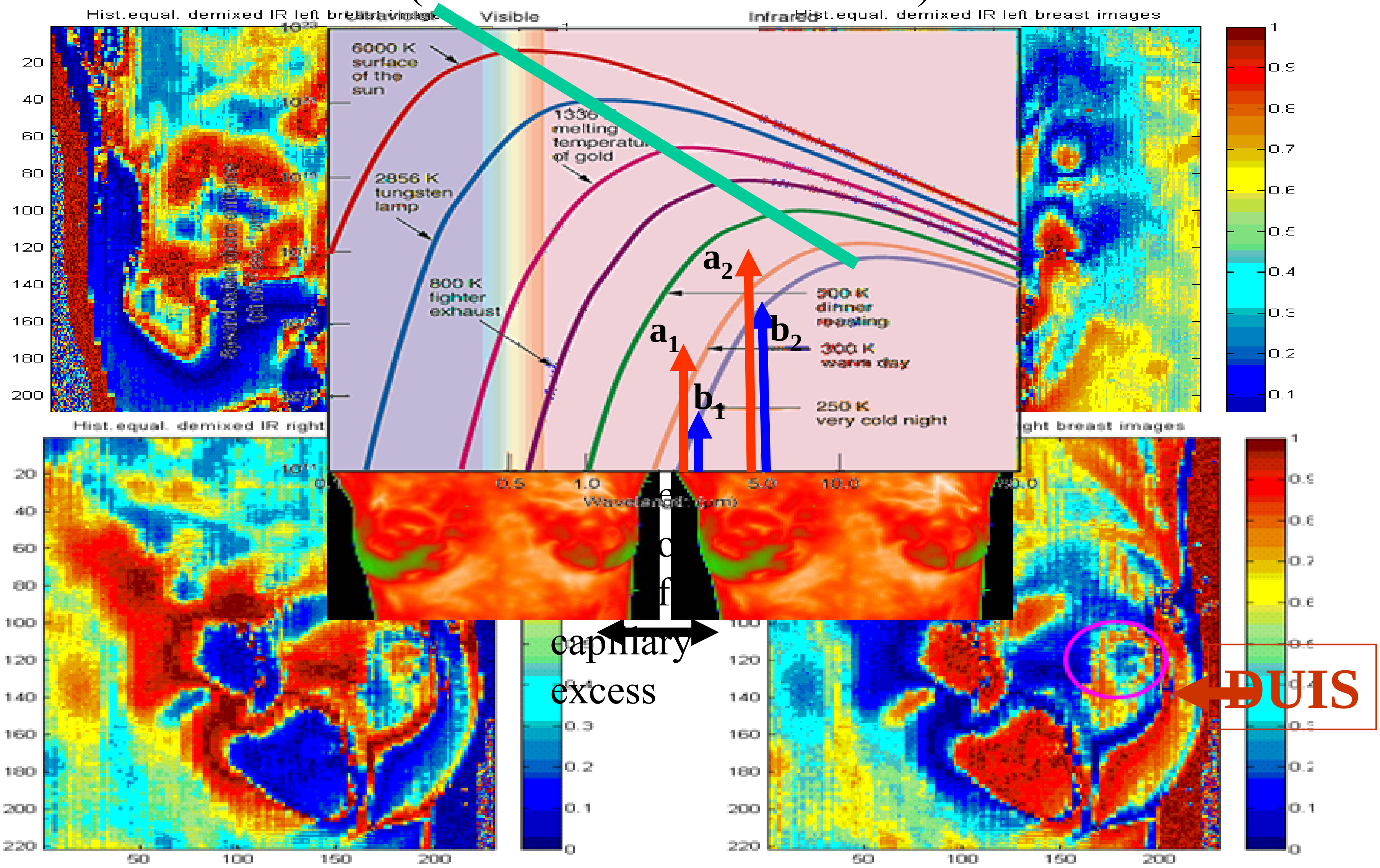


Long IR

Two cameras passive imaging track risky patients without radiation hazard and waiting

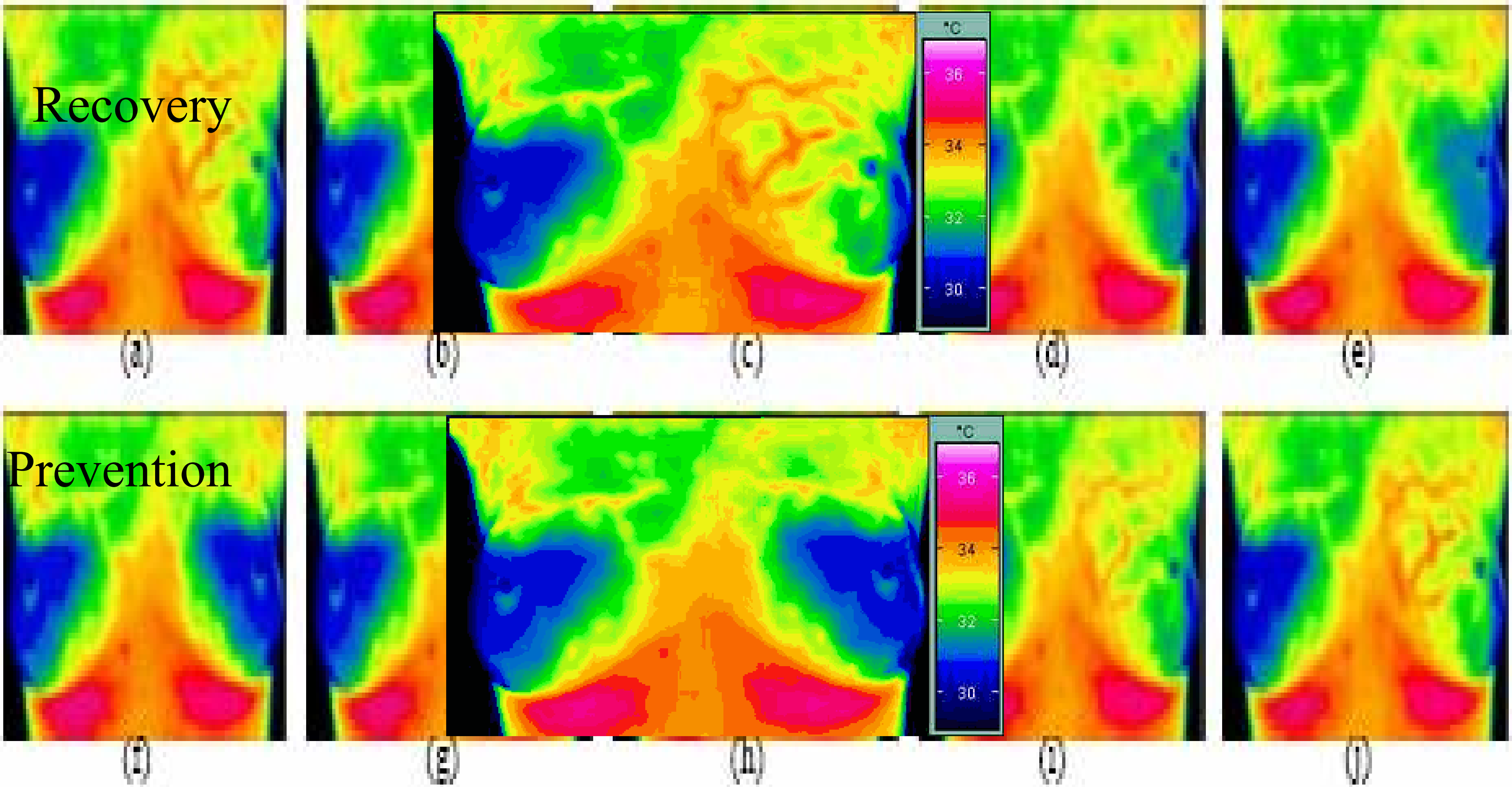
Reference: Christine Gorman," Rethinking Breast Cancer", ***Time Magazine***, p.p.50-58, Feb 8, 2002

Appendix Impact Area: Unsupervised Learning based Gibb's Spontaneity Principle of minimum Isothermal Helmholtz Free Energy (ONR Press Release 2000)

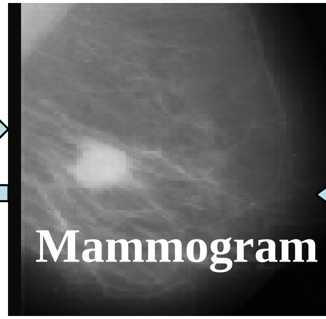


A2. Harold Szu 2007

St. John Paul Pope Reverse Protocol based on "***Randomized Double-Blind 50-50 Mixture of Sick & Non-sick Patients with Referee Tracking the Recovery History***" (IP pending) with physician ground truth for ***Rec.Op.Char. (PD vs. FAR)*** saving the usual search about 100K patients for early diagnoses in a time reversal way



Vatican Protocol:Due to religious concerns, we augmented the necessary practice of Mammogram, biopsy, augmented with passive high-resolution IR Spectrograms in time.



In-situ Analyses

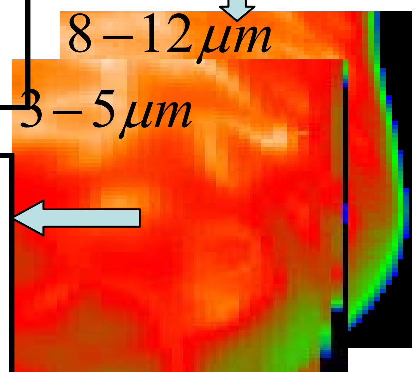
Localized Malign Tumor(s) ?

Yes, but what stage & degree?

Oncology Medical Treatments (Chemotherapy, etc.)

Now, Satellite- graded high-resolution high-cost (at $20K x 2) liquid nitrogen, two IR dual-color cameras

to build for every household the 1-D CNT-NanoPixels, no liquid Nitrogen,a digital IR spectrograms Camera at $1K



Smart Algorithm: single-pixel unsupervised Blind Decision aids

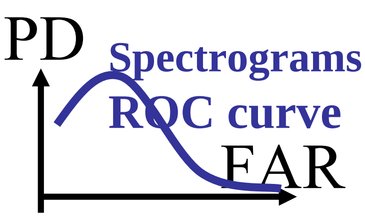
Percentage degree per pixels of maligns

Fully recovered?

NO

YES

Animated spectrograms movie, treatment history, playing backwards reveals seeding of angiogenesis effect; About 50~100 cases of successes give

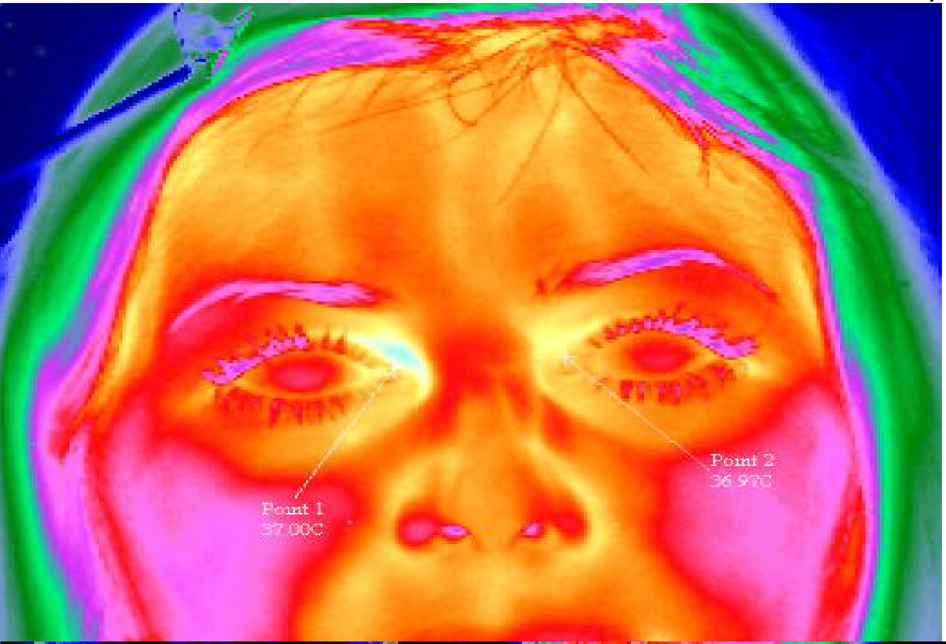


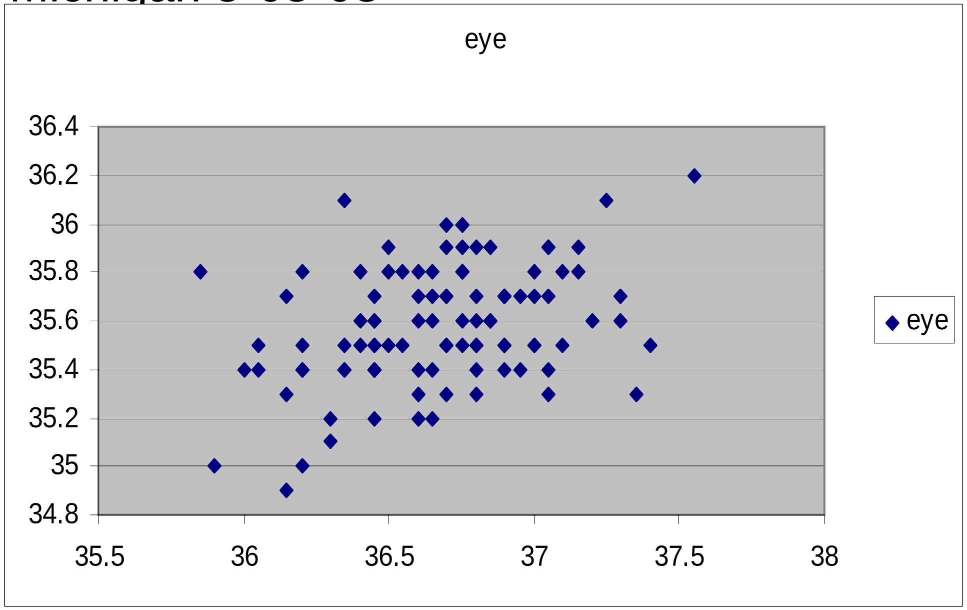
Passive Dual-color infrared Spectrograms tracking system

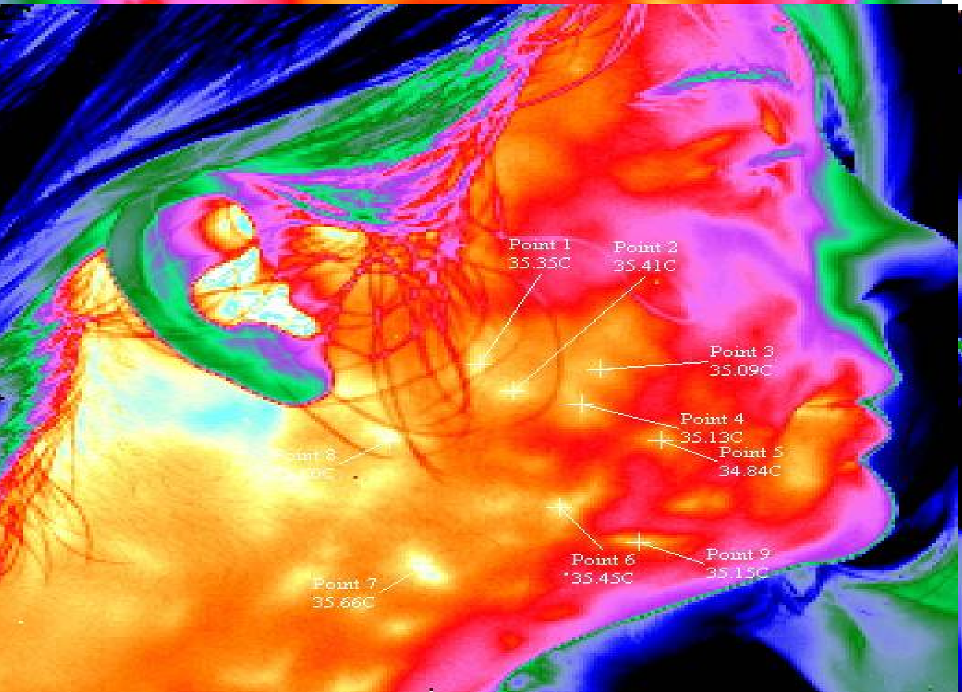
A3. Harold Szu 2007

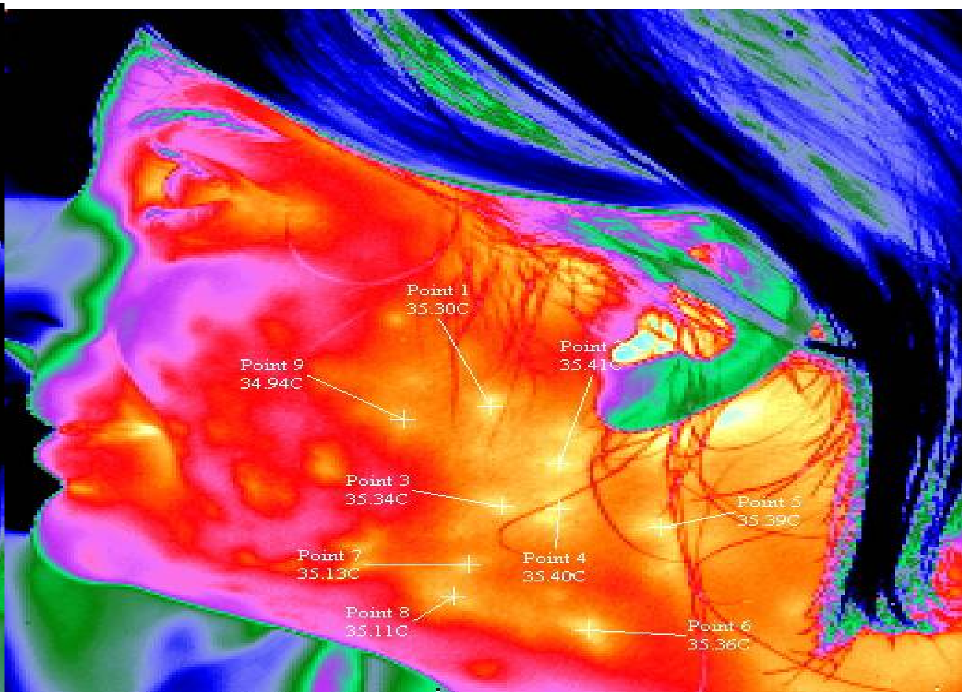
EXPERTS RE-DEFINE SARS AIRPORT THERMOGRAPHY SCREENING

Detroit/Toronto , Michigan 5-05-03

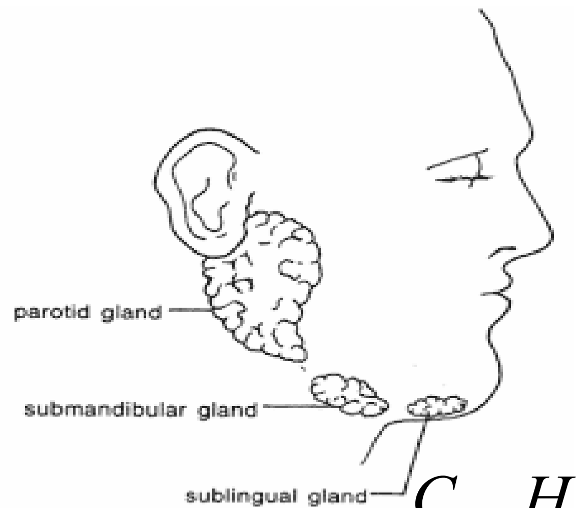


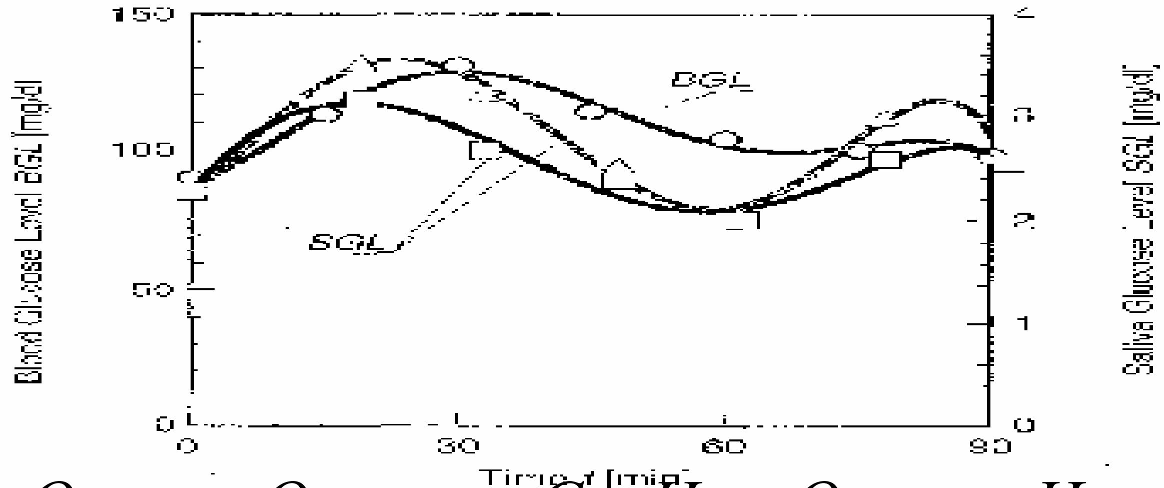




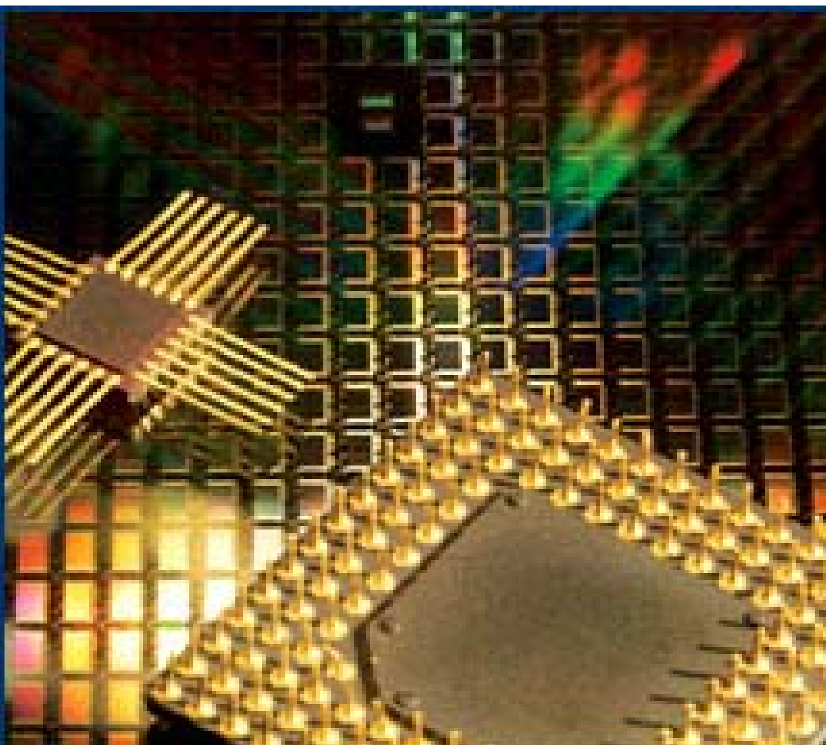


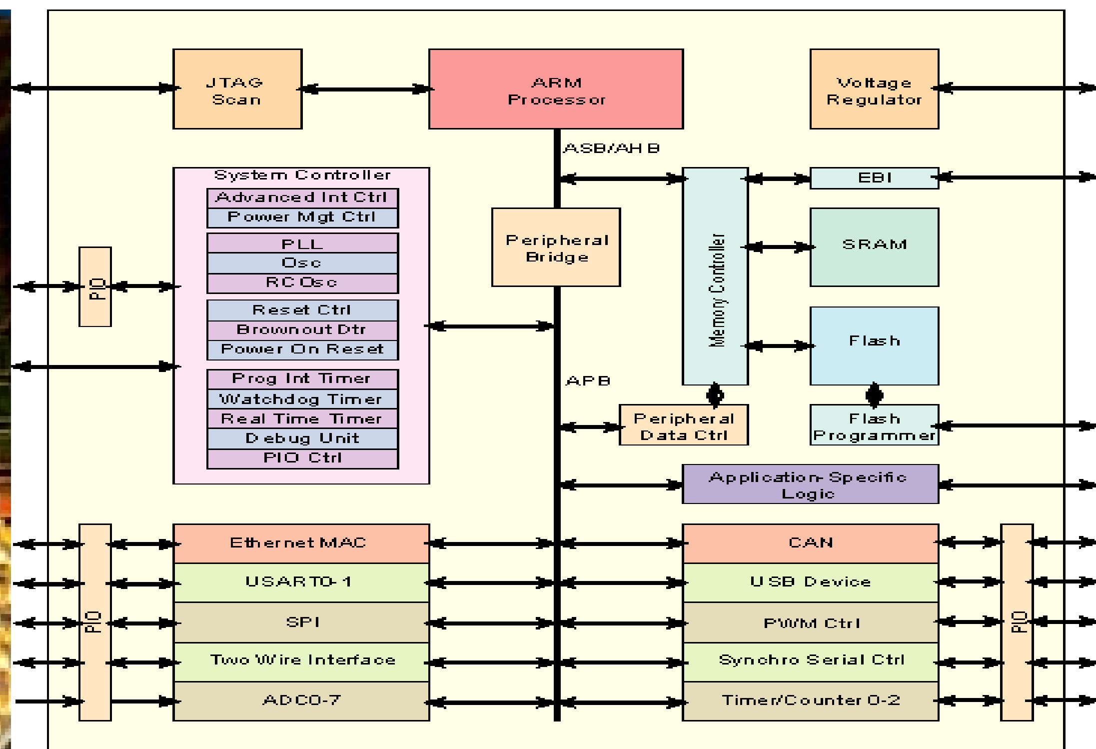
Saliva Test SOC augmenting Blood Test





$$C_6H_{12}O_6 + O_2 \rightarrow C_6H_{12}O_6 + H_2O_2$$





Multidisciplinary Computing Intelligence Magazine (ICI) serve

•the niche of credible wellness knowledge benchmarked with timely R/D know-how for the aging population home care needs.

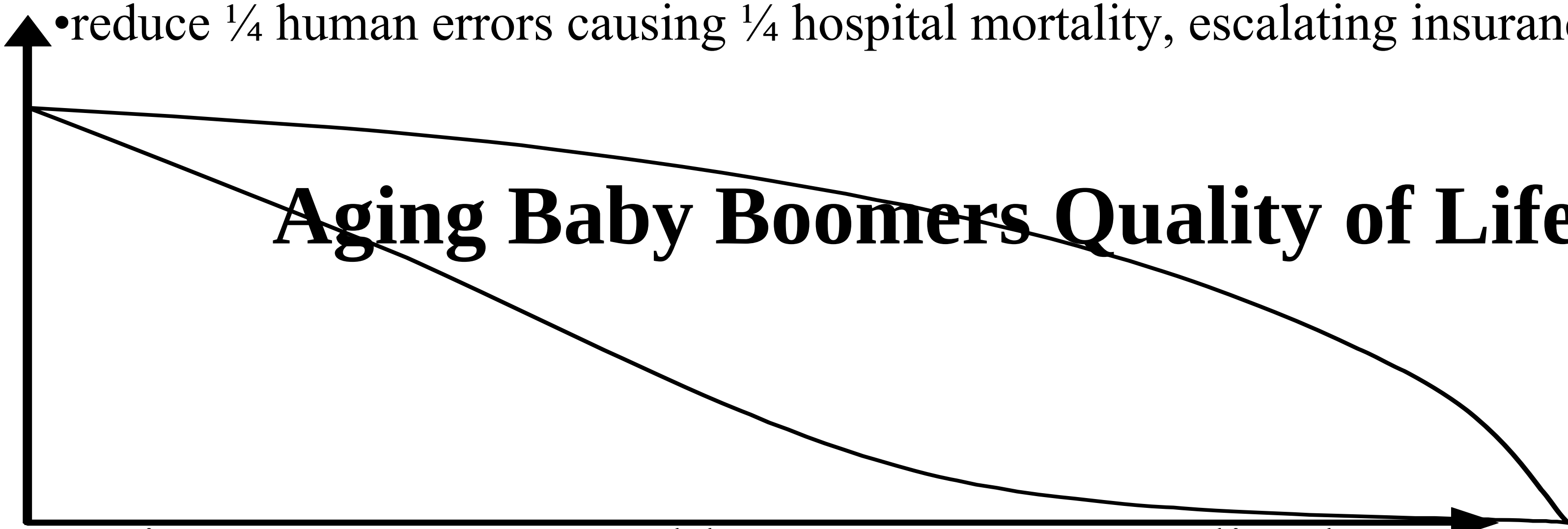
•Measuring daily the ***“Wellness Baseline Profiling (WBP)”*** can save healthcare cost by improving the quality of life of ¼ US Population (78M aging baby boomers) costing ¼ GDP with

•the earliest options for diseases, dementia screening, diagnoses, and treatments.

Quality of Life Assurance:

•***WBP*** gives better ***Individual Diagnosis Aids (IDA)*** for “***a stitch in time saves nine***” which helps

•reduce ¼ human errors causing ¼ hospital mortality, escalating insurance cost.



Retirement at 65 years old; average man lived to 78 years and woman at 80

Generalized Brain Info Theory beyond Shannon

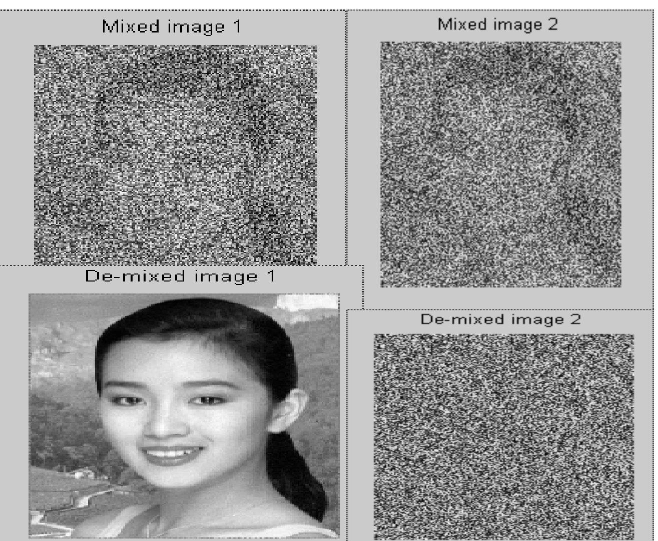
US Patent Trade Office Publication Szu et al. 2004

(i) Two observations turn out to be the necessary & sufficient conditions for **learning without teacher**, **breakthrough of** **unsupervised ANNs**

(ii) Apply **physics to physiology to derive** **Brain Info Theory generalizing brain info theory with proven unsupervised learning capability**.

(1) Why "Power of 2" for inputs? But outputs are singular sensor!

(2) Why constant "37°C" ? **"Homeostasis Learning" Szu, Rus. Acad. Sci. 1999**

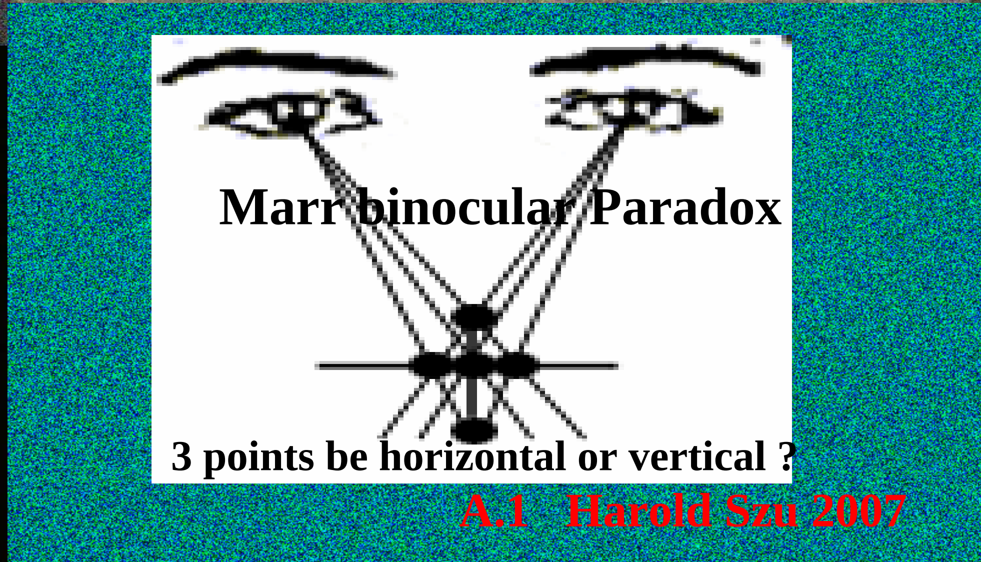


$$Vector \quad \vec{X}(t) = [A(t)?]\vec{S}(t)?$$

$$2x5 = 10 \ D \ redundancy \ \ reduction$$

"Agree, the signal; disagree, the noise" requires no teacher and instantaneous.

Impact Area: Why do we have pairs, e.g. two eyes, etc. ten sensors for five sense inputs? **A reason is "*agree, the info;disagree, the clutter*" at the constant 37^{o}C brain temperature for a soft-decision of agreement for instantaneous response for the survival of fittest species (Unsupervised Learning)**



Theory of Brain Info Theory

- Shannon info theory borrowed from physics, Boltzmann statistical mechanics
- Shannon theory is valid for a closed equilibrium system, e.g. Bell Telephone.
- Brain is open dynamic system that has a constant ***temperature*** & deals with ***pairs of sensory inputs*** 5x2=10 dimensional vector time series, overly redundant data
- From these two observation alone, we have derived from the physics the
- **(1) Unsupervised Neural Network Learning Hebb Rule,**
- **(2) Neuron Sigmoid logic and predict**
- **(3) Ion Current** directly involved with the ***synaptic junctions*** in the learning, mediated perhaps by house-keeping ***glial cells***.

Generalized Info Theory beyond Shannon

(i) Two observations turn out to be the necessary & sufficient conditions for **learning without teacher**, **breakthrough of unsupervised ANNs Szu** 1997

(ii) Apply **physics to physiology to derive Brain Info Theory generalizing Shannon info theory**.

(1) Why "Power of 2"?

$$Vector \qquad \overset{\omega}{X}(t) = [A(t)?]\,\overset{\rho}{S}(t)\,?$$

$$2\,x\,5 = 10 \quad D \quad redundancy \qquad reduction$$

(2) Why "37^{o}C" ?

But the output has one unique sensor!

"Homeostasis Learning" Szu, 1999 Rus. Acad. Sci.

ANN involves dendrite signal pre-conditioning

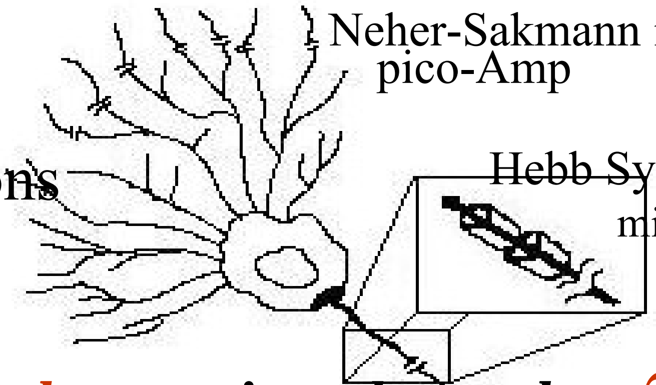
Isothermal Free Energy ***$H = E - T_o S$*** ***Helmholtz***

"Gasoline does PdV work after the heat waste"

Where a homeostasis body ***$T_o = 37^oC$,***

$$K_B T_{room} = \frac{1}{40} eV = Nhf$$

Thermal Energy could support 10B neurons



Entropy ***$S = - NK_B \Sigma_i s'_i \log s'_i$***

Local Lagrange ***Constraint,*** ***1991Neher-Sakmann*** ***ion channels*** ***μ, λ.***

$$E = \vec{\mu}_{10^{-12}\ Amp} \left\{ [W]_{10^{-3}\ Volt} \vec{X} - \vec{S} \right\} = \vec{\lambda} \left\{ [A] \vec{S} - \vec{X} \right\}$$

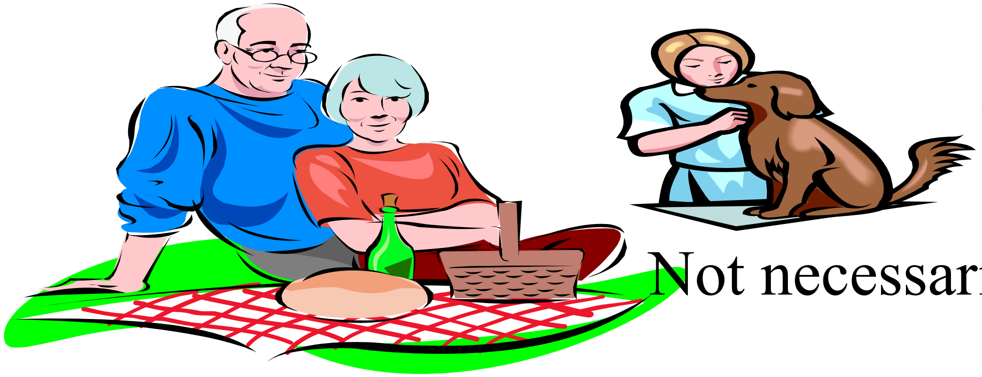
Lagrange multipliers were conjectured to play the role of ion current & housekeeping glial cells in neurobiology for the unsupervised internal tutor.

Pair Inputs need unsupervised redundancy reduction to feature Supervision is only for the label of feature subspace

Childhood learning experience of a "dog" begins with highly redundant inputs- playful experience of eyes, ears, etc. 5 senses10 dimensional – not by Fisher neighborhood separation

5-sense 10-D data $X_{dog}(t)$: effortlessly (fusion [W] X_{dog})→ $S_{dog}(t)$

S_{dog} = (s_1,s_2, unsupervised features) $_{dog}$ fuzzy dog label by supervision

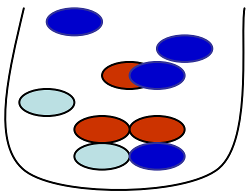


Not necessarily by separating dogs from Cats

Parents call out "a dog", a label of fuzzy linguistic nature of which the vector span of the subspace is self discovered by unsupervised manner, for no parents ever defined the dog.

Shannon's Theory came from a closed equilibrium

Independent & identical color balls ***(N=R+G+B)*** of single bucket have the ***a priori*** combinatorial state space ***W=N!/(R!G!B!)***



Boltzmann's Tomb Stone $S = K_B \, \mathrm{Log} W$

of which Shannon formula follows

$$S = -NK_B \, \Sigma'_i \, s_i \log s_i > 0$$

where the prime indicates $\Sigma_i s_i = 1$, ***the minus sign for log*** $s_i < 0$ ***as*** $s_i < 1$.

Proof: ***If N=R+G+B, then 1=R/N+G/N+B/N and use is made of Stirling formula:*** $\log N! = \log N(N-1)(N-2)\ldots. = N \log N - N$

Then,

$$S/N = K_B \, \mathrm{Log}(\, N!/R!B!G!)$$

$$= K_B[(R+G+B)\log N - \underline{N} - (R\log R - \underline{R} + B\log B - \underline{B} + G\log G - \underline{G})]$$

$$= K_B \, [-R/\log(R/N) - B \log(B/N) - G \log(G/N)]$$

$$= -NK_B \, [s_1 \log s_1 + s_2 \log s_2 + s_3 \log s_3] = -NK_B \, \Sigma'_i \, s_i \log s_i.$$

QED

2. Generalized Shannon Info Theory to Brain Info Theory

by minimum of Helmholtz thermodynamic free energy,

Min. $H = E - T_o S$ **(Max. by Shannon under** $T_o = 37^oC$**)**

Szu assumes analytic I/O $E = \mu([W]X - S) + \lambda([W]X - S)^2$ **+Taylor**

1997 Theorem for unsupervised sensory fusion: If info energy E is analytic in info I/O , (data X & feature S), then the necessary & sufficient conditions of unsupervised learning in Dude-Hart unlabelled data classifier sense are:(1) An intelligent brain is kept at constant temperature, e.g. human 37^{o}C

Unsupervised Learning

(2) All input sensors are Smart Pairs : "Power of Pairs In, Garbage Out"

Information is kept within memory

1. IEEE Press 2004 "Comp. Intel" Ch.16 Szu Unsupervised Learning ANN,

2. Shanghai Sci Ed Publ. 2003 Szu & Zhang" Intel Image Proc. Blind Sources Sep.

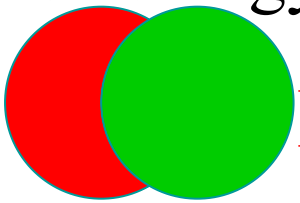
Mathematical Ill-posed Inverse Problem "Power of Pairs"

give vector time series X(t) having numerous feature S(t) decompositions upon which one chooses the most probable equilibrium answer imposed by min.Helmholtz free energy

Guess what were hidden real positive energy sources 3 &5 ?

e.g. $2 \times 5 + 1 \times 3 = 13$

$1 \times 5 + 3 \times 3 = 14$



Power of Pairs

Given two resulting numbers 13 & 14 ***Given data X***,

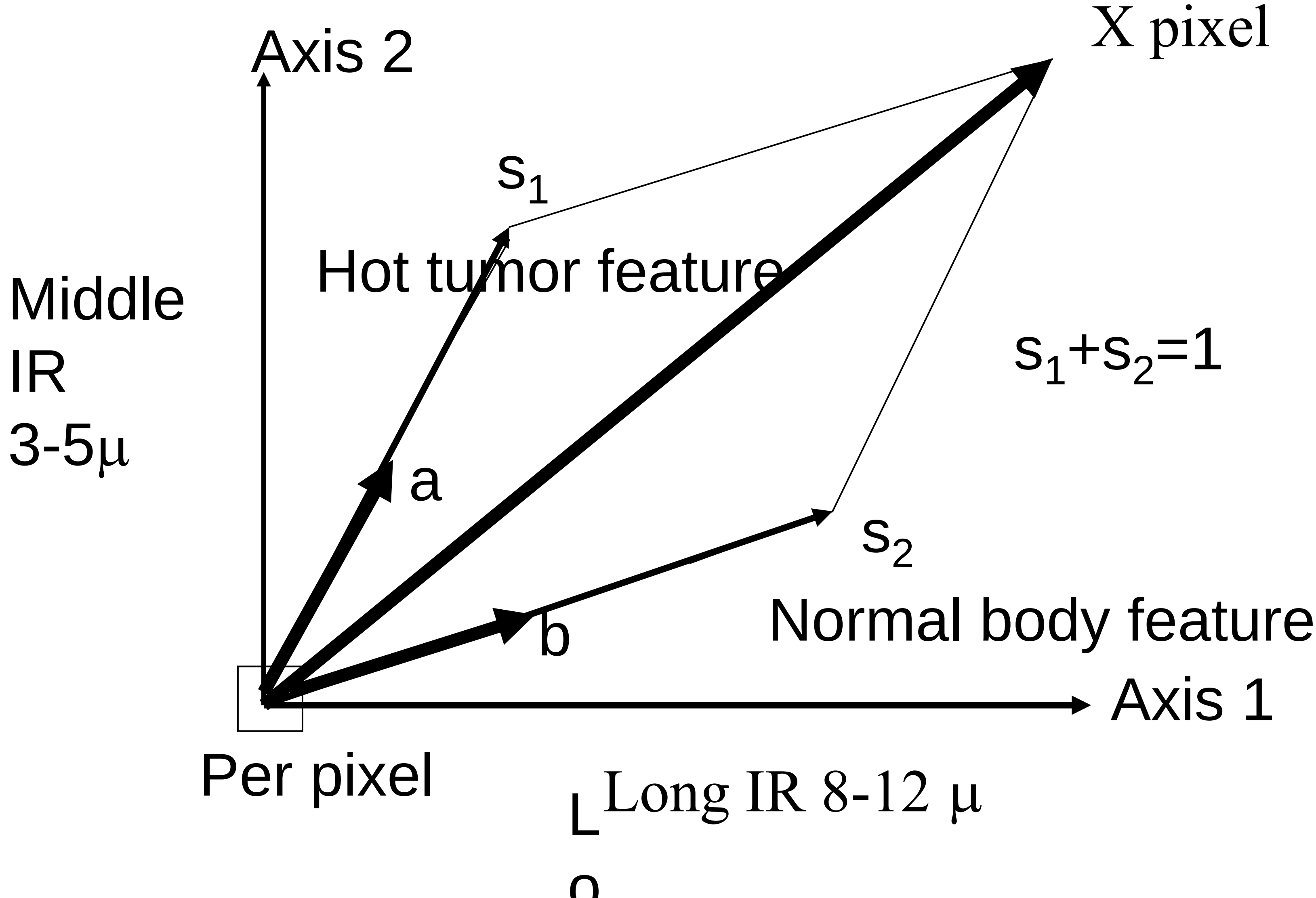
find both unknown mixing matrix & sources [A?]S?

$$\begin{bmatrix} 13 \\ 14 \end{bmatrix} = 5\begin{pmatrix} 2 \\ 1 \end{pmatrix} + 3\begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}\begin{pmatrix} 5 \\ 3 \end{pmatrix}; \vec{X} = [A]\vec{S} = s_1\vec{a} + s_2\vec{b};$$

we can always normalize the data X

for unknown unit feature vectors $\vec{a}, \vec{b}$

$$knowing\ component\ s_1,\ s_2\ follows$$
$$uniquely\ vector\ features\ [\vec{a}\ \vec{b}] \equiv [A]$$

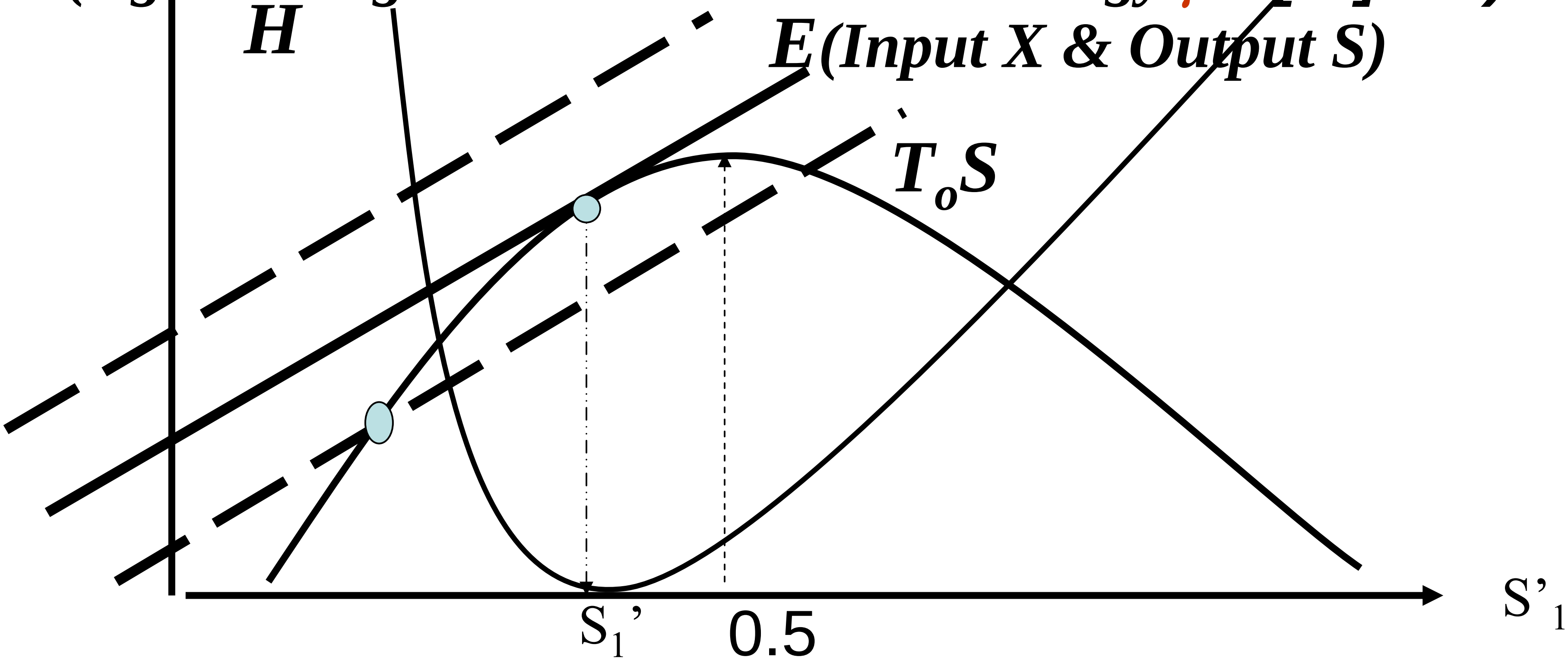


Graphical Proof of Uniqueness of sources

We postulate a linear information I/O energy

$$E = \mu\{[W]\mathbf{X}-\mathbf{S}\} = \lambda\{[A]\mathbf{S}-\mathbf{X}\}$$

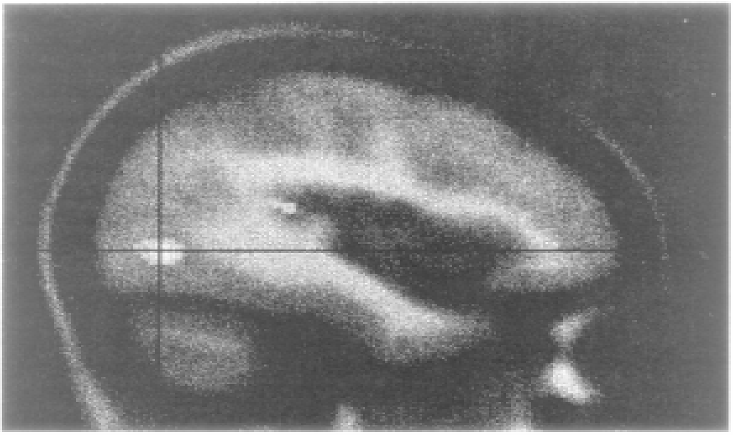
(e.g. which generalizes LMS error energy $\mu = [W]\mathbf{X}-\mathbf{S}$)



Local Maximum Entropy for Most Probable Solution

$s_2' = 1 - s_1'$

Homeostasis Learning derives naturally Hebb & Sigmoid rules for isothermal resource sharing among glia cells & neurons

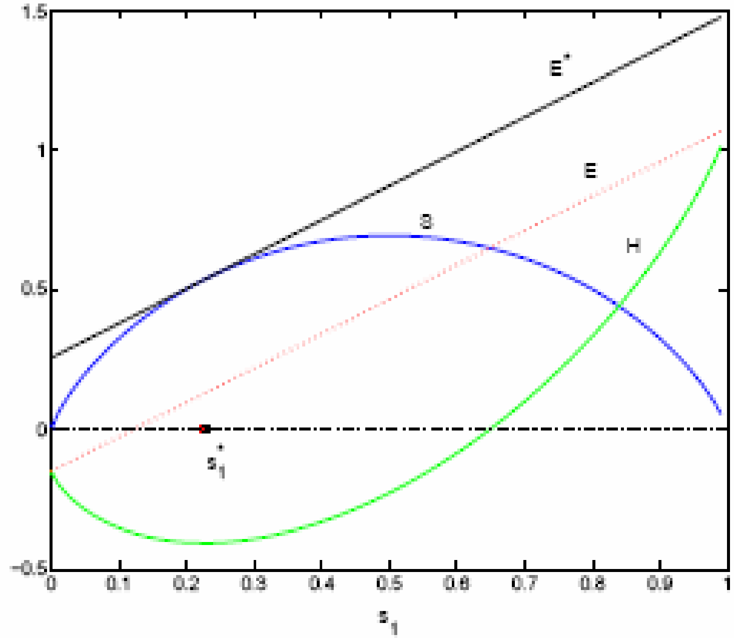


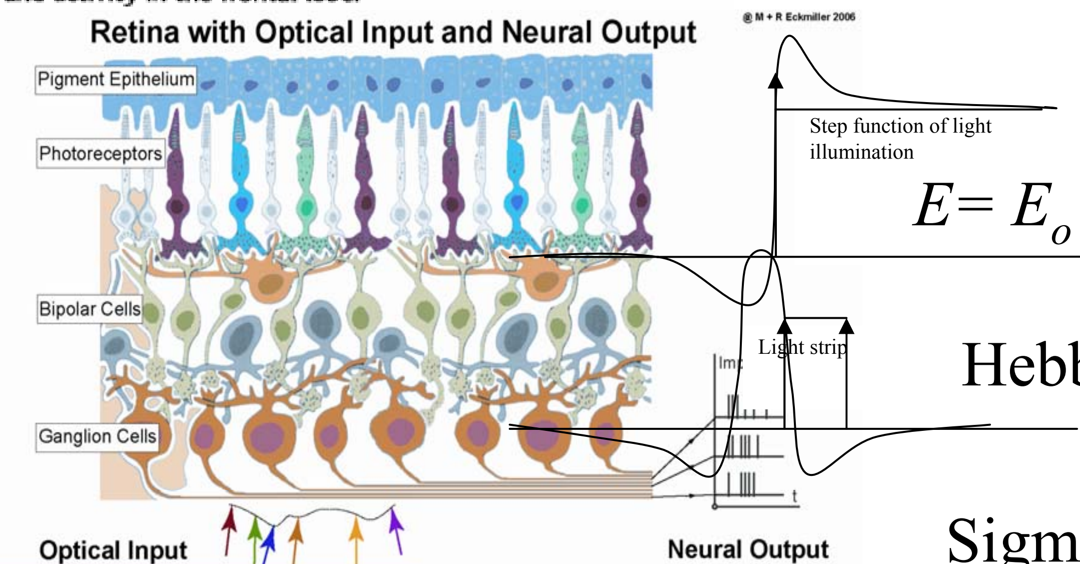
Busy brains. Unlike low scorers on a spatial test, high scorers show a correlation between activity in a visual area (cross-hairs) and activity in the frontal lobe.

$$H \equiv E - T_o\ S;$$

$$H = 0;$$

$$\frac{dH}{ds_1} = 0$$

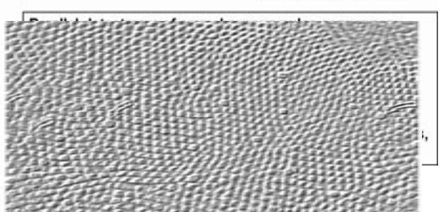




$$E = E_o + \boldsymbol{\mu}\ \boldsymbol{([W]X - S) + second\ order}$$

Hebb $$\frac{\partial W_{ij}}{\partial t} = -\frac{\partial H}{\partial W_{ij}} \Rightarrow \Delta W_{ij} \approx \mu_i X_j$$

Sigmoid $$\frac{\partial H}{\partial s_i} = 0 \ \Rightarrow \sigma(\mu_i) = s_i$$



p.2 Harold Szu 2007

Closed Equilibrium *Theorem of Maximum A Priori Entropy:*
m-Equal-Partition Law

Given a-priori entropy of m independent identical parts

$$S = -NK_B \Sigma_{i=1,m} s_i \log s_i$$
$$= -NK_B \Sigma_{i=1,m} s_i \log s_i + NK_B (\mu_o+1)(\Sigma_{i=1,m} s_i - 1)$$

Proof: ***One can not conduct partial differentiation as all components are coupled by the unit norm, unless we remove it by Lagrange multiplier (μ_o+1)***

$$S = -NK_B \sum_{i=1,m} s_i \log s_i + NK_B (\mu_o+1)(\sum_{i=1,m} s_i - 1)$$

$$\frac{dS}{ds_j} = \sum_k \frac{\partial S}{\partial s_k}\frac{ds_k}{ds_j} \rightarrow \left.\frac{\partial S}{\partial s_j}\right|_{const.k} = NK_B(\log s_j + 1) + NK_B(\mu_o+1) = 0$$

$s_j = \exp(\mu_o)$ ***to determine Lagrange multiplier by the constraint value***

$$\sum_{j=1,m} s_j = \sum_{j=1,m} \exp(\mu_o) = m\exp(\mu_o) = 1$$

$$s_j = \exp(\mu_o) = \frac{1}{m} \qquad Q.E.D.$$

Open dynamic equilibrium by the minimum of Helmholtz free energy $H=E-T_oS$ generalized the equal-partition law to Sigmoid Law

$$H([W],\vec{S}) = E - T_0 S$$

$$= \vec{\mu}^T\left[[W]\vec{X} - N\vec{S}\right] + NK_BT_0\sum_{i=1}^{m} s_i \log s_i - NK_BT_0(\mu_0+1)\left(\sum_{i=1}^{m} s_i - 1\right);$$

$$N = Norm\,\vec{X}$$

$$Equilibrium\ at\ \frac{\partial H}{\partial s_j} = 0 \Rightarrow$$

$$\vec{s}_j == \exp\left(\frac{\mu_j}{K_BT_0} + \mu_0\right) = \frac{1}{1 + \sum_{\substack{i=1\\ i\neq j}}^{m} \exp\left(\frac{1}{K_BT_0}(\mu_i - \mu_j)\right)} = \sigma(\vec{\mu}_j)$$

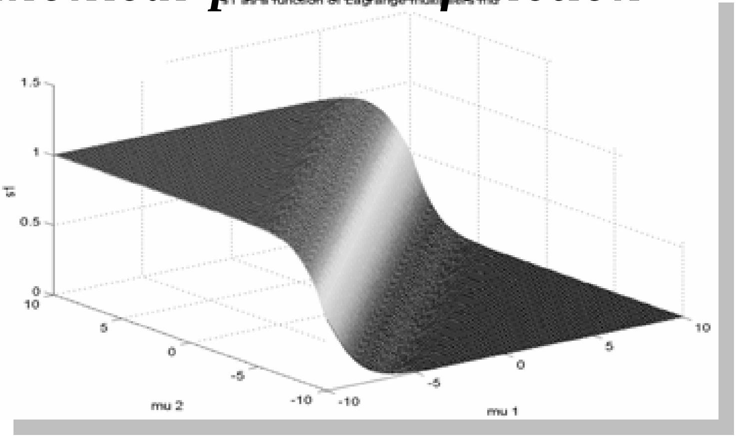
$$where\ use\ is\ made\ of\ the\ constraint\ value\ \sum_{j=1}^{m} s_j = 1$$

$$\Rightarrow \exp(-\mu_0) = \sum_{j=1}^{m} \exp\left(\frac{\mu_j}{K_BT_0}\right) \equiv Z\ \ Canonical\ partition\ function$$

One can verify the free energy

as displacement with constraint force

$$\frac{\partial \mu_0}{\partial \mu_j} = -s_j$$



Biomedical Spontaneity Gibb's

Theorem 1 Minimum Helmholtz free Energy determines a stationary function of excited abnormal or malign state:

Given isothermal equilibrium system at the minimum Helmholtz free energy defined the departure from an individual baseline level satisfies a canonical ensemble average behavior, in terms of analytic input & output data ***X*** *and* ***S****. Then, the departure of the baseline according to minimum thermodynamics free energy one obtains*

H(isothermal free energy) = (info energy)E - (local temperature)T_o (Shannon entropy)S

$$E = E_o + \sum_{i=1,2} \frac{\partial E}{\partial s_i^{(o)}} (s_i - s_i^{(o)}) = E_o + \sum_{i=1,2} \mu_i (\sum_{i,j=1,2} [W_{ij}] X_j - S_i)$$

$$X(t) = a_o [A(t)?] S_o(t) \quad E = E_o + [\mu_1 \ \ \mu_2] \begin{bmatrix} (W_1^T, X) - s_1 \\ (W_2^T, X) - s_2 \end{bmatrix}$$

$$= E_o + \mu_1 (W_1^T, X) + \mu_2 (W_2^T, X) - \mu_2 + (\mu_2 - \mu_1) s_1$$

$$H \equiv E - T_o S ;$$

$$H = 0 ;$$

$$\frac{dH}{ds_1} = 0$$

$$s_1^* = 1 - \exp(-\frac{E_o^*}{K_B T_o})$$

Two State Exact Solution: Proof

In order to determine the correct zero-order formulation for the baseline, one need to derive the relationship between the true solution E_o^* (indicated by the superscript *) and the true unknown source s^*_1. Shannon using the unit-sum-rule entropy formulae:

$$S = -s_1 \log s_1 - (1-s_1)\log(1-s_1)$$

The first order slope of energy tangent line at each point is obtained by calculating the derivative of the entropy *S* with respect to

$$\frac{dS}{ds_1} = -\log s_1 - \frac{s_1}{s_1} + \frac{1-s_1}{1-s_1} + \log(1-s_1) = \log(1-s_1) - \log s_1$$

Since at equilibrium $H=E-T_oS=0$, then $E=T_oS$, and the slope of the energy line equals to the slope of the entropy line. If we know the(i) intercept E_o from the immediately previous baseline measurement, the energy line can be uniquely determined

$$E/K_B = slope \cdot s_1 + intercept =$$

$$T_o \frac{dS}{ds_1} \cdot s_1 + E_o^* = T_o\{\log(1-s_1) - \log s_1\}s_1 + E_o^*/K_B$$

Then, substituting the aforementioned entropy formula *S* and equaling it to the derived energy E formulae (because(ii) ***H=0*** implying $\boldsymbol{E=T_oK_BS}$), we have obtained $S=E/K_BT_o$

$$S = -s_1 \log s_1 - (1-s_1)\log(1-s_1) =$$

$$\{\log(1-s_1) - \log s_1\}s_1 + E_o^*/K_BT_o = E/K_BT_o$$

$$E_o^*/K_BT_o = -\log(1-s_1^*) \qquad s_1^* = 1 - \exp\left(-\frac{E_o^*}{K_BT_o}\right)$$

Nonlinear Lagrange multipliers becomes arbitrary, λ to $\lambda \bullet C(s)=\lambda \bullet 0=0$ when correct constraint is approached. The perturbation $C(s) \equiv WX - s = 0$ $\Delta X_i = \sum_{j=1}^{M} \frac{\partial X_i}{\partial \lambda_j} \Delta \lambda_j = \sum_{j=1}^{M} J_{ij} \Delta \lambda_j$ where Jocobian matrix $C(s)=(WX-s)$ becomes the error covariance. $J_{ij} = (W_{i\alpha} X_\alpha X_\beta W_{\beta j} - s_i s_j)$

Then the inverse determines the Lagrange multiplier $\Delta\lambda_j = J^{-1}{}_{ji} \Delta X_i$ requiring the singular value decomposition (SVD) of the Order O(M^2) computation. To reduce O(M^2) to O(M) as follows.

Kuhn and Tucker developed in 1951 Lagrange multiplier method, of which **Karush** introduced a quadratic penalty function called augmented Lagrange multipliers. We rewrite the Helmholtz free energy $H=E-T_oS$ in the following Lagrange:

$$\underset{s}{Max}\ L_{c^k}(s,\ \lambda,\ \lambda_o) = f(s) + \lambda \bullet C(s) + \frac{1}{2c^k} \| C \|^2$$

$$vector \quad \cos t\ C(s) = WX - s$$

$$Negative \quad Entropy \ -T_o S \equiv f(s) \equiv T_o K_B \sum_{j=1}^{M} s_j \log s_j + (\lambda_o - T_o K_B)(\sum_{j=1}^{M} s_j - 1)$$

The necessary condition of KKT is defined as follows: $L(s^*,\ W,\ \lambda^*,\ c^k \to \infty) = f(s^*) + \lambda^* \bullet C(s^*)$

(i) $\lambda^* \bullet C(s^*) = 0$

(ii) $\nabla L(s^*,\ \lambda^*) = \nabla f(s^*) + \lambda^* \bullet \nabla C(s^*) = 0$

(iii) $\nabla L(s^*,\ \lambda^*) = \nabla f(s^*) + (\lambda + \frac{1}{c^k} C) \bullet \nabla C(s^*) = 0$

From (ii) and (iii) follows an order O(M) $\lambda^* = \lambda + \frac{1}{c^k} C$ decoupled linear equations:

References:

H. W. Kuhn and A.W. Tucker "Augmented Lagrange Multipliers," In: "Nonlinear Program," the 2[Nd] Berkeley Symposium Mathematics and Probability (ed. J. Neyman, Pub. UC Berkeley, 1951).

D. P. Bertsekas, "Constraint Optimization and Lagrange Multiplier Methods," (LANCELOT Algorithm) Academic 1982

If the breast tissue takes ΔE energy from its surroundings, the entropy change of S_r will be $\Delta S_r = -\Delta E/T_0$, and the total entropy change is

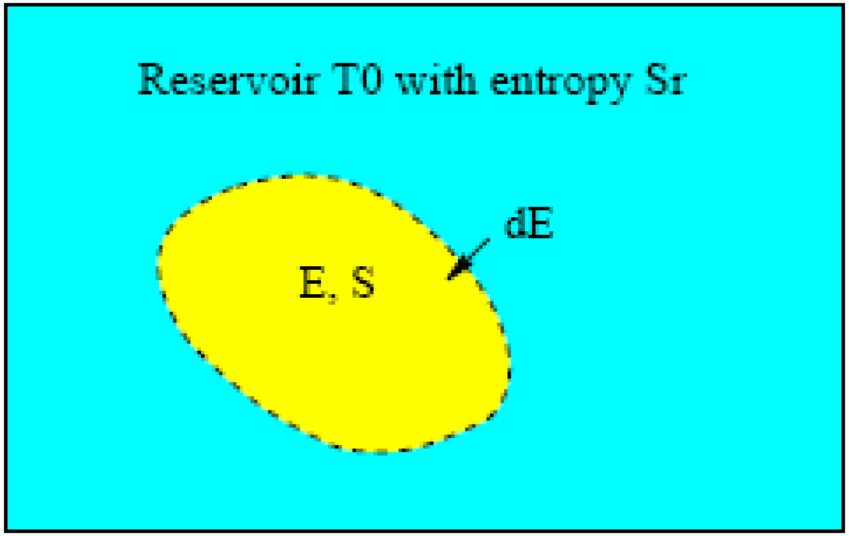


Figure 3. The system diagram.

$$S_t = S_r + S$$

$$\Delta S_t = \Delta S_r + \Delta S = -(\Delta E - T_0 \Delta S)/T_0 = -\Delta H/T_0$$

where $\Delta H = \Delta E - T_0 \Delta S$ is the change of Helmholtz free energy of the breast tissue system.

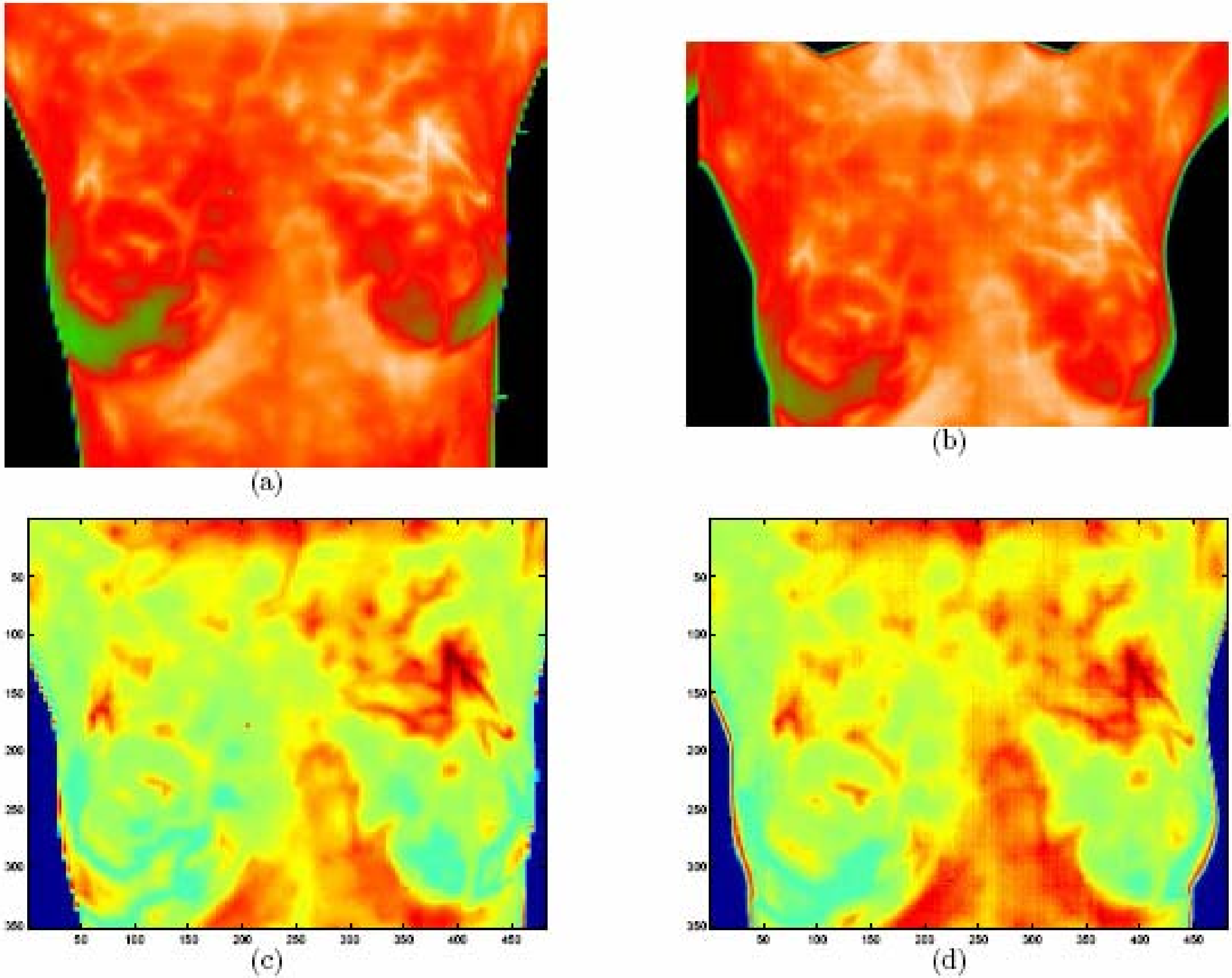


Figure 6. The original and registered IR images. (a) Original middle wavelength IR (b) Original long wavelength IR (c) Registered middle wavelength IR (d) Registered long wavelength IR.

Plank Law & Data Vectors

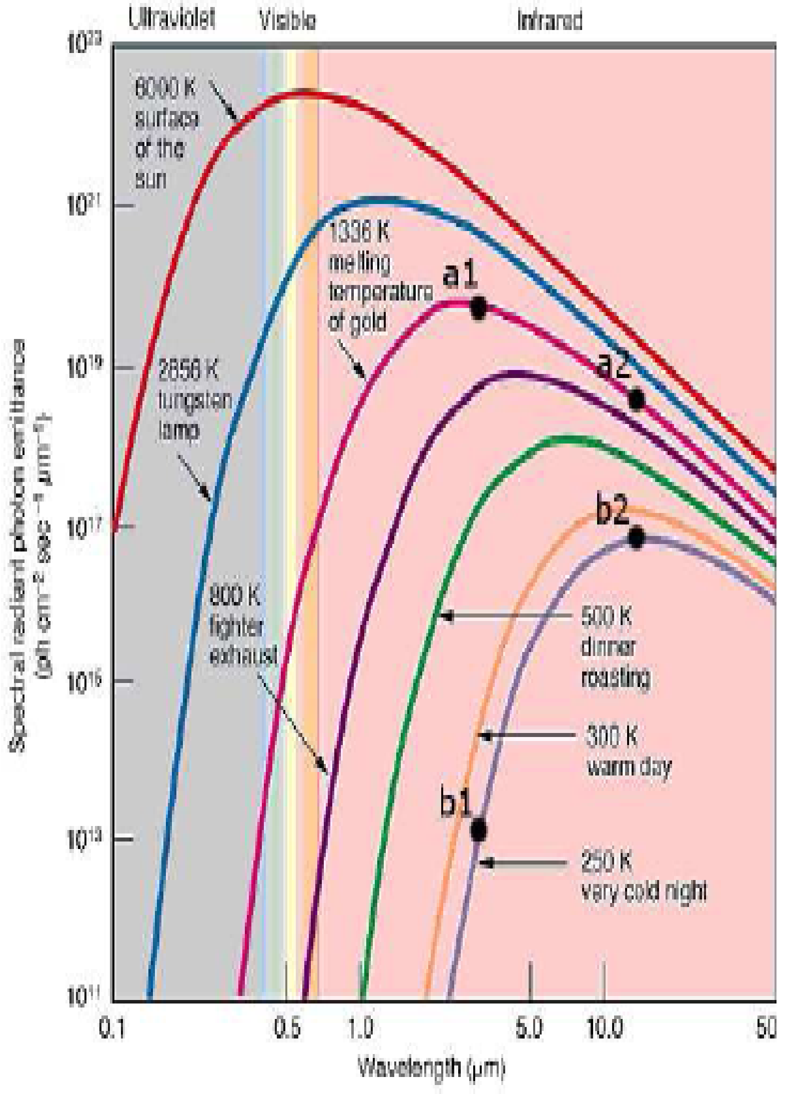


Figure 1. The Plank radiation law.

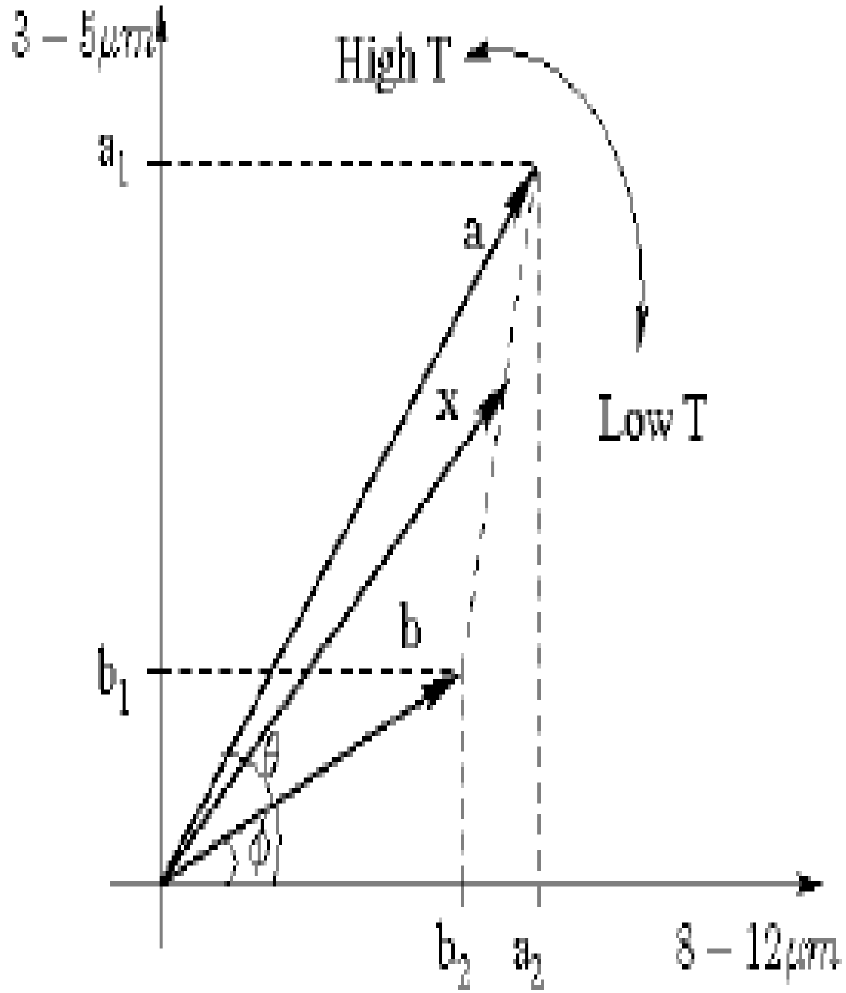


Figure 2. The vector diagram of linear mixture model.

Unsupervised Learning

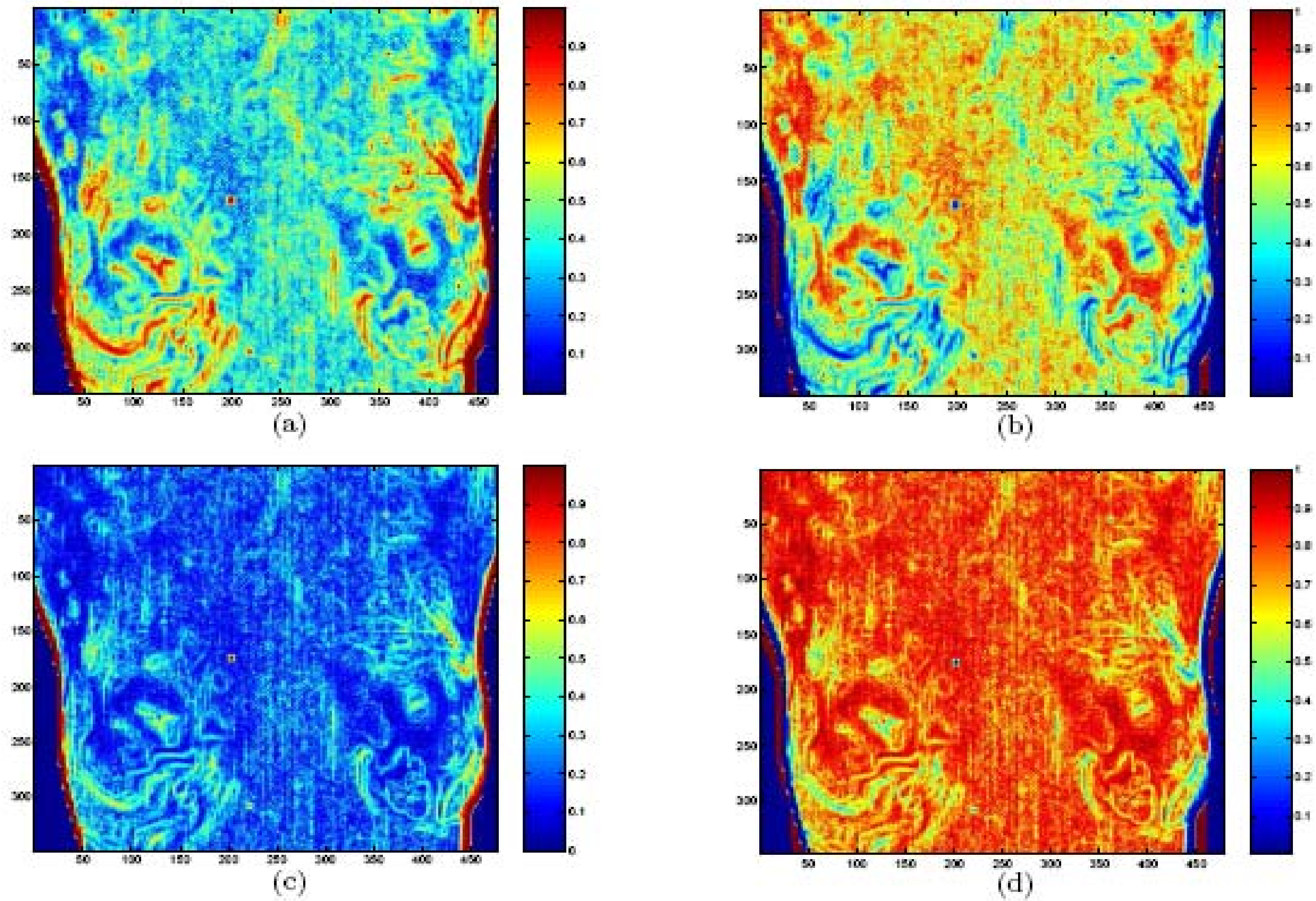


Figure 9. Unsupervised classification results using the deterministic single pixel BSS algorithm. (a) Source 1 (fixed neighborhood) (b) Source 2 (fixed neighborhood) (c) Source 1 (adaptive neighborhood) (d) Source 2 (adaptive neighborhood)

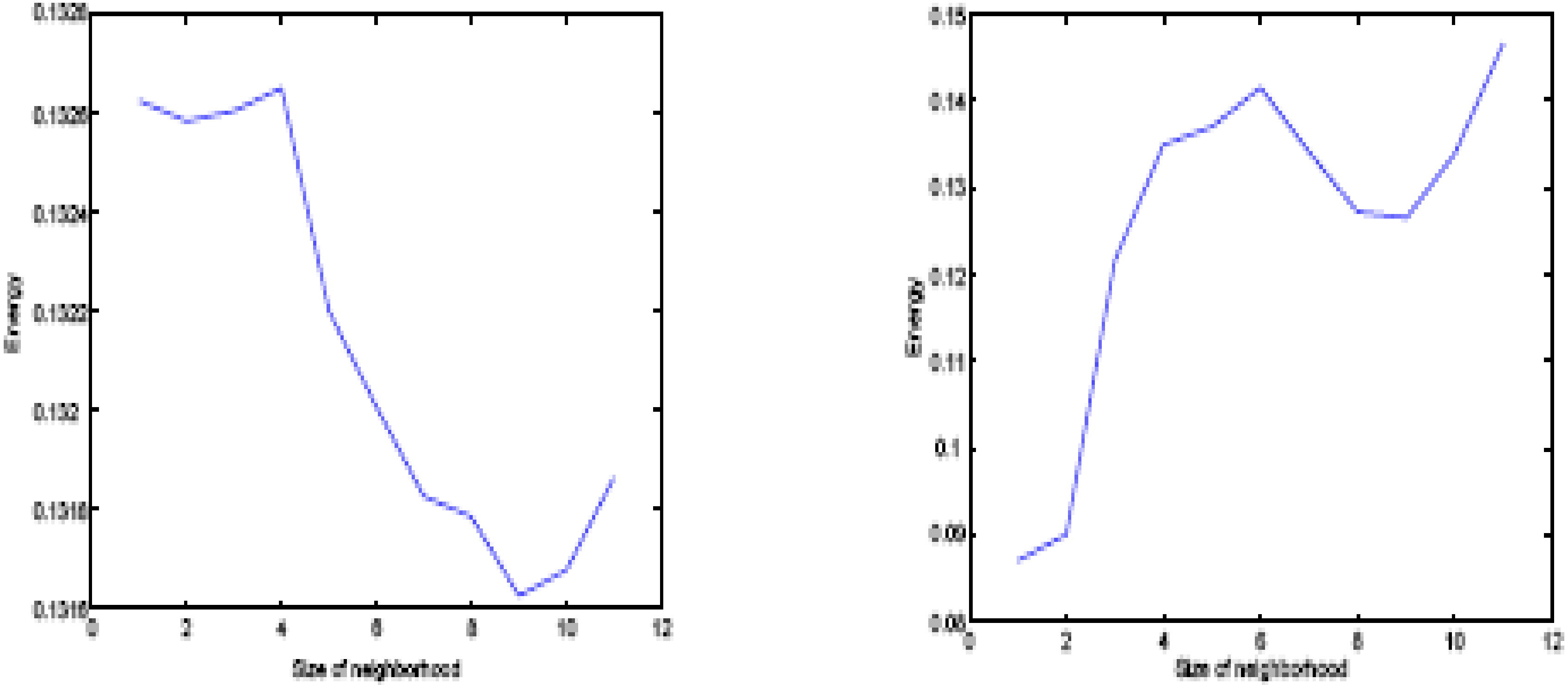


Figure 5. The ground state energy as a function of neighborhood size.

1^{St} & 2^{nd} Lagrange Constraint NN

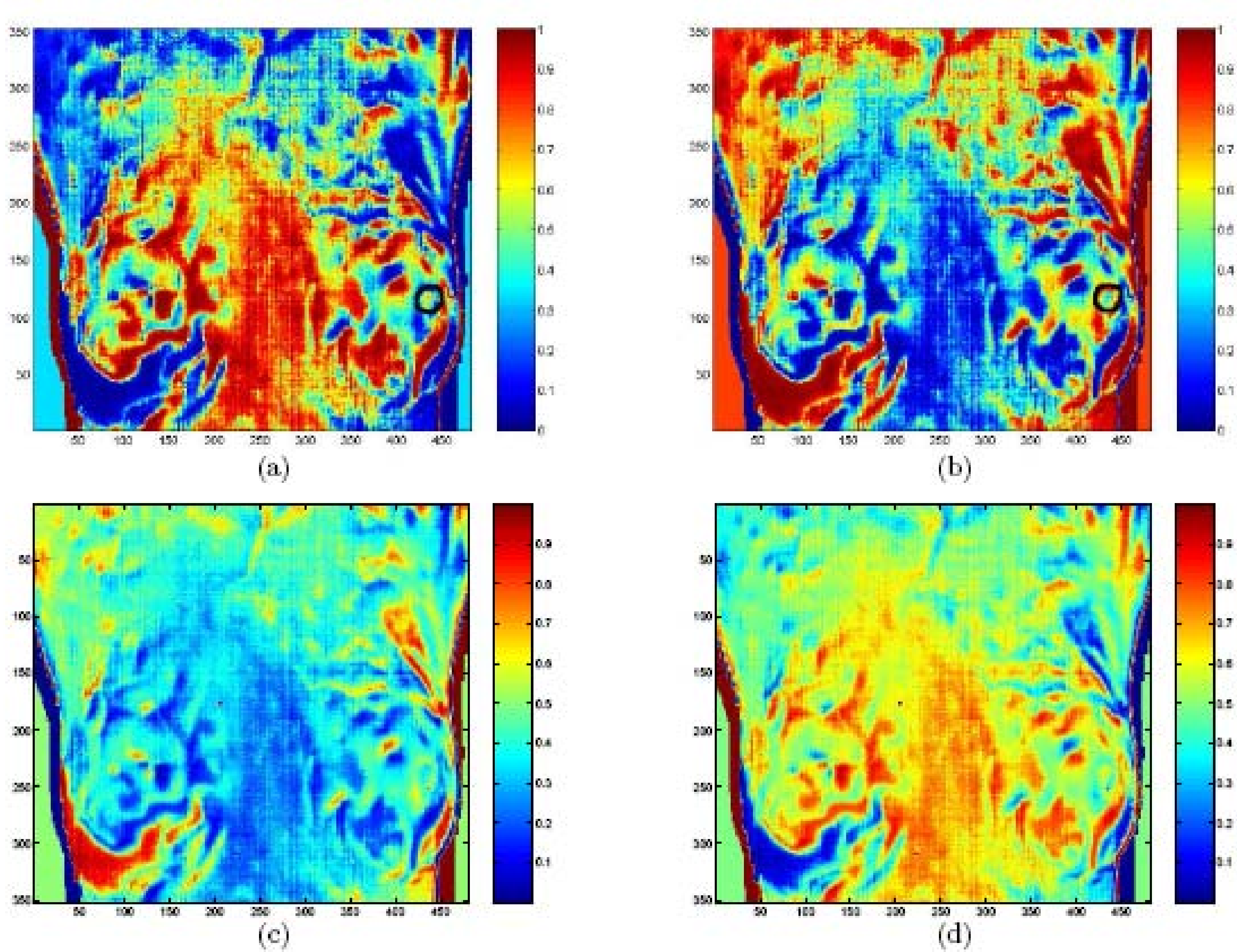
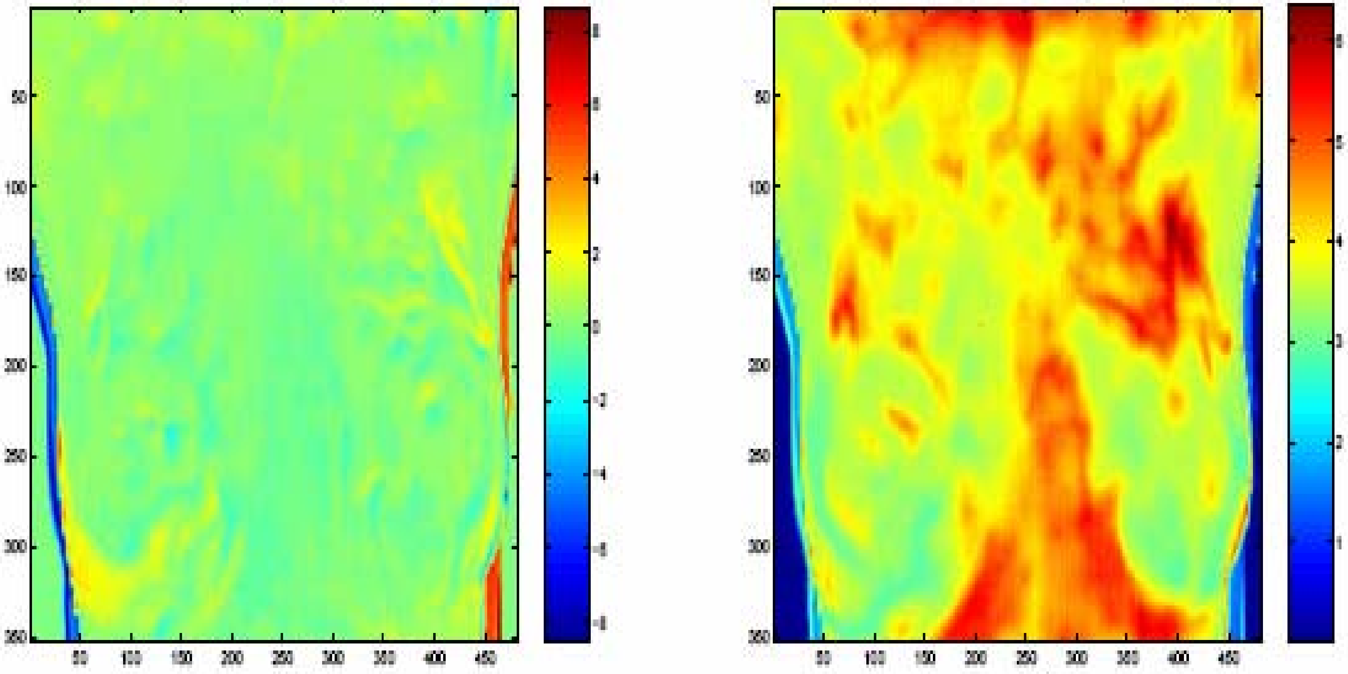


Figure 7. Unsupervised classification results using the first- and second-order LCNN. The marked area on the left breast indicates the existence of abnormal tissue. (a) Source 1 (1st-order) (b) Source 2 (1st-order) (c) Source 1 (2nd-order) (d) Source 2 (2nd-order)

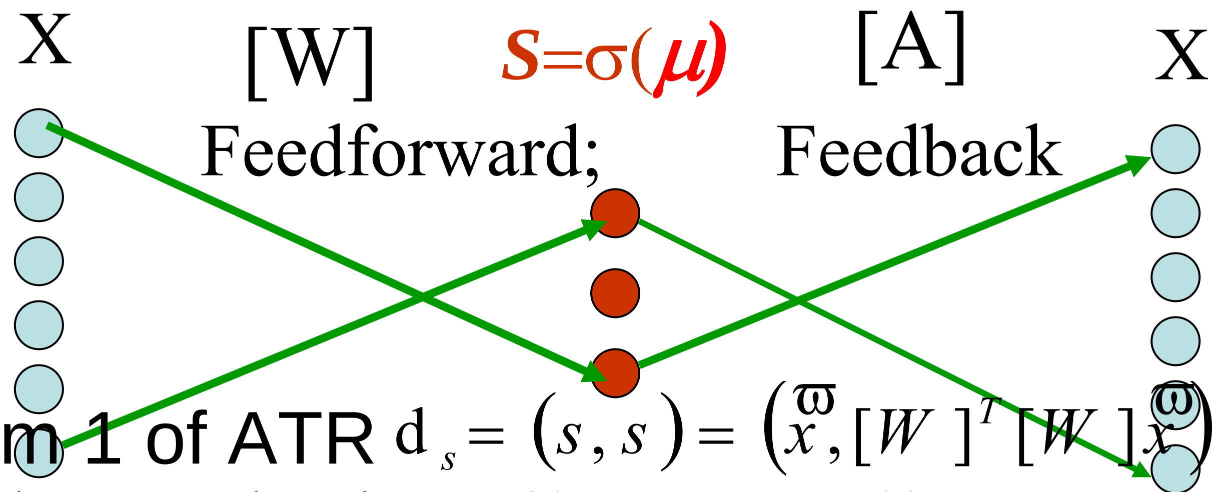
Fast ICA



(a) (b)

Figure 8. Uncovered sources using the FastICA, which are still mixtures of normal and abnormal tissues. No abnormal tissue is identified. (a) Source 1 (b) Source 2

Unsupervised Learning Based on Thermodynamic Equilibrium by Min. Free Energy H=E-TS; E=μ([W]X-S)= λ(X-[A]S) 1st order LMS error energy



Theorem 1 of ATR $d_s = (s, s) = (\vec{x}, [W]^T [W] \vec{x}) = d_{\vec{x}}$

- Uncertainty Reduction S(t)?=[W?]X(t);
- Associative recall X(t)=[A?]S(t)?
- Theorem 2 Unsupervised Hebb info distance

full rank learning (d[W]/dt)= - (dH/d[W]) $[W^T][W]$ = μX $[W^T][W]$

- Theorem 3 Generalized Equal Partition Sigmoid logic

$$\sigma(\mu) = S = [W]X$$

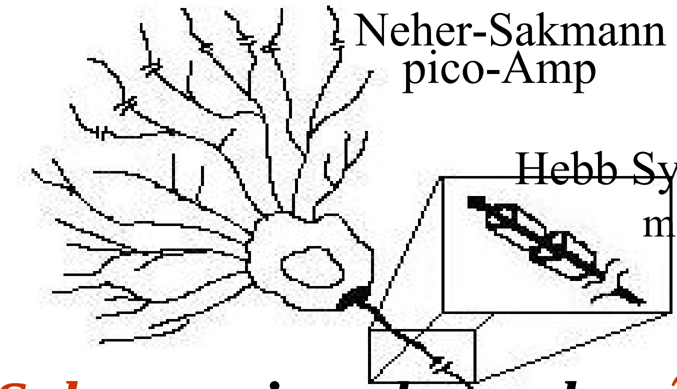
ANN involves dendrite signal pre-conditioning

Isothermal Free Energy $H = E - T_o S$ *Helmholtz*

"Gasoline does PdV work after the heat waste"

Where a homeostasis body $T_o = 37^oC$,

$$K_B T_{room} = \frac{1}{40} eV$$



Entropy $S = - NK_B \Sigma_i s'_i \log s'_i$

Local Lagrange Constraint,1991Neher-Sakmann ion channels μ, λ.

AM uses 10,000 neurons

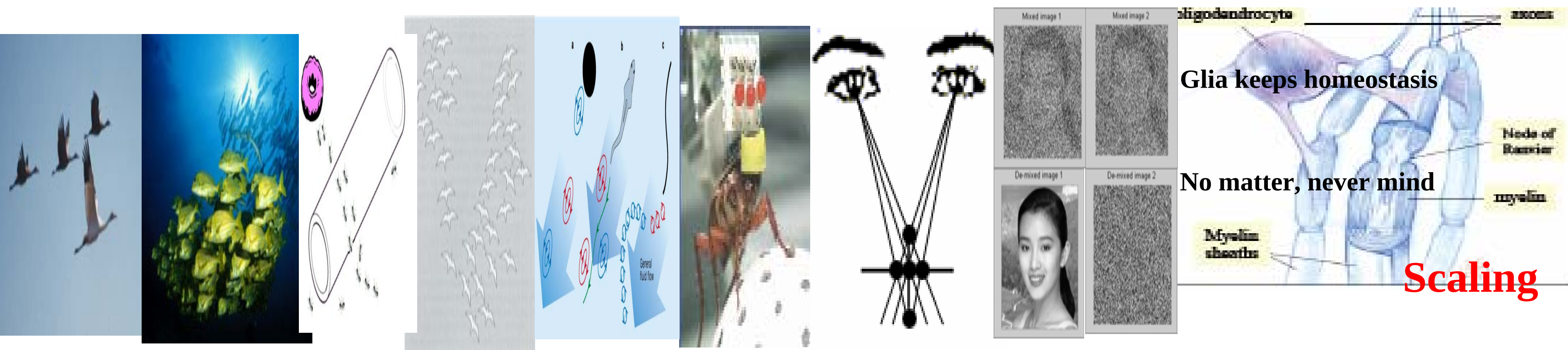
$$E = \vec{\mu}_{10^{-12}\ Amp} \left\{ [W]_{10^{-3}\ Volt} \vec{X} - \vec{S} \right\} = \vec{\lambda} \left\{ [A] \vec{S} - \vec{X} \right\}$$

Lagrange multipliers were conjectured to play the role of ion current & housekeeping glial cells in neurobiology for the unsupervised internal tutor.

Claustrum is the legacy Crick left behind:

Crick & Koch identified a thin sheet of gray matter resides in two way parallel connection below the cortex, computing feelings, seeing, hearing, language, deciding what to do. They said: ***"Claustrum is a conductor of consciousness like an orchestra that binds all senses". Philos. Trans. Royal Soc. Lond, 2005***

From Consciousness to mind from mind to body vice versa in society



From Physics To Physiology estimates 20 B of Neurons

H. Szu 2007

•If one takes Crick-Koch conscious model---***Claustrum*** seriously, then the ubiquitous binding of bursting 30Hz firing rates is the wakening of consciousness.

•What would be the size of neurons & house-keeping glia cells in our brain kept at 37°C that can support effortlessly? It turns out to be tenths billions.

•Owing to Boltzmann and Einstein formulae, our equilibrium brain size may be estimated as follows: ((1/40) $eV = K_B T_o$ at 300°K room temperature);

•*$E = Nhf = K_B T_o$,ranging hundredth to tenth* ***billions glias & neurons*** *corresponding to* ***30Hz and 60 Hz*** *firing rates respectively for the ground state of mindlessness.*

$$E = N\,hf = N\,6.67 \times 10^{-34} f \ \text{Joule}$$

$$\approx K_B T_o = 1.38 \times 10^{-23} \times 310 \ \text{Joule}$$

$$\text{Thus}, \quad N \le 2 \times 10^{11}, \ \text{if} \ f = 30 \ Hz$$

$$N \le 10^{11}, \ \text{if} \ f = 60 \ Hz$$

$$N \cong 6 \times 10^{10}, \ \text{if} \ f = 100 \ HZ$$

•One shall tabulate the sizes of brains with respect to the temperatures kept by **warm blood animal kingdom** for neuroethological and development mechanism.

•Far from equilibrium life forms and evolves. Math-Physics waits for new method.

•Equilibrium physics value provides local insight and boundary condition.

Long Term: physiology & physics elucidate, exploit sub-consciousness (sublimated awareness) **to early perception** (reasoning, memory) **& self learning** (supervised, unsupervised, decision, action) for ***swarm robotic intelligence*.**

Mid Term: **Substantiate** house-keeping ***glia cells "μ_j" functionality*** as the ***missing half of Einstein's brain***, maintaining synaptic junctions providing. "(1) Why do we have pairs of sensors? (2) Why constant temperature brains? What is the function and behavior of glia at individual synapses? Should the interactions among glia cells and neurons follow the Gibbs spontaneity:min. $H = E_0 + \sum_i \mu_i \{[W_{ij}] X_j(t) - s_i(t)\} - T_o S$ where $T_o = 37^0C$; S= entropy; E=I/O energy? We derived well known

(i) Hebb Rule: $$\frac{\partial W_{ij}}{\partial t} = -\frac{\partial H}{\partial W_{ij}} \Rightarrow \Delta W_{ij} \approx \mu_i X_j$$

(ii) Sigmoid Rule: $$\frac{\partial H}{\partial s_i} = 0 \Rightarrow \sigma(\mu_i) = s_i$$

Near Term: To build test-bed of **biomimetic pairs of biosensors** and to demonstrate a homeostasis **Unsupervised Learning** capability for signal and noise enhancement, & feature extraction for **Biomedical Wellness (BMW)** engineering.

High-Impact: **Interdisciplinary Computing Intelligence (ICI)** serve the wellness of aging population measuring daily the "Wellness Baseline Profiling (WBP)" to save healthcare cost by improving the quality of life of ¼ US Population (78M aging baby boomers) costing ¼ GDP with the earliest possible diseases and dementia.

Review of ANN

We derive ANN from top-down & bottom-up observations:

(1) Bottom-up:vector time series input from pairs sensors

(2) Top-down: isothermal equilibrium condition.

We obtain three important results as follows:

(i))We shall first review Boltzmann entropy that leads to Shannon formula by Stirling then generalize it.

(ii) The equilibrium equal partition law, which becomes for brain open equilibrium to be sigmoid logic without assuming it.

(iii) We assume Lagrange Constraint m to be Neher-Sakmann

pico-Ampere ion channels 1991.

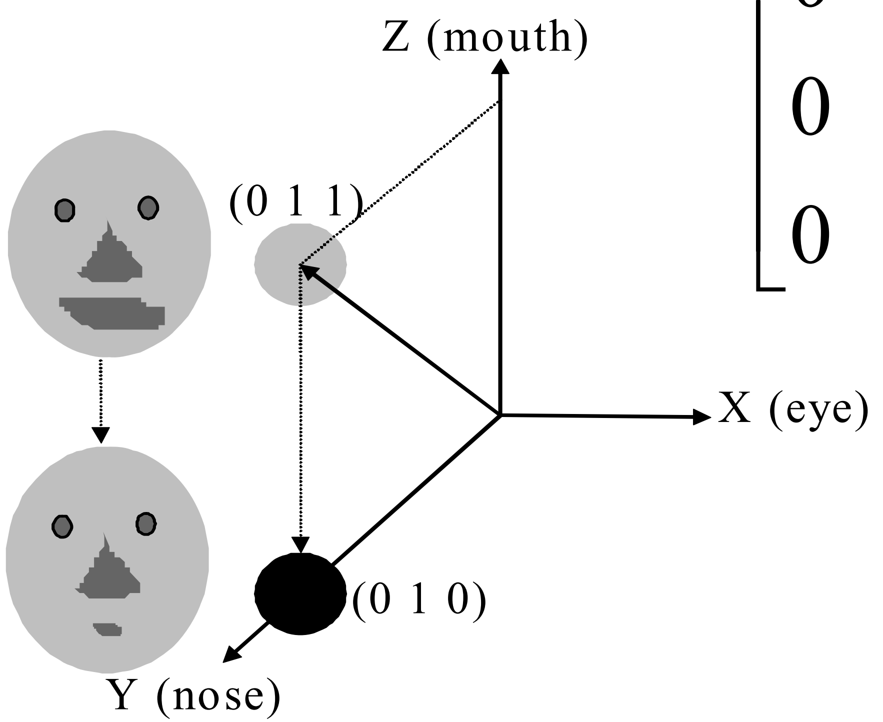
(iv) We derive unsupervised Hebb rule from free energy min.

(v) Constraint optimization based on Cauchy fast simulated annealing at cooling schedule $T=T_o/(1+t)$, $t>0$.(vs.$T_o/\log t$)

Fault Tolerant AM Read by 2D Matrix Vector Inner Product

- [AM][Vector]=

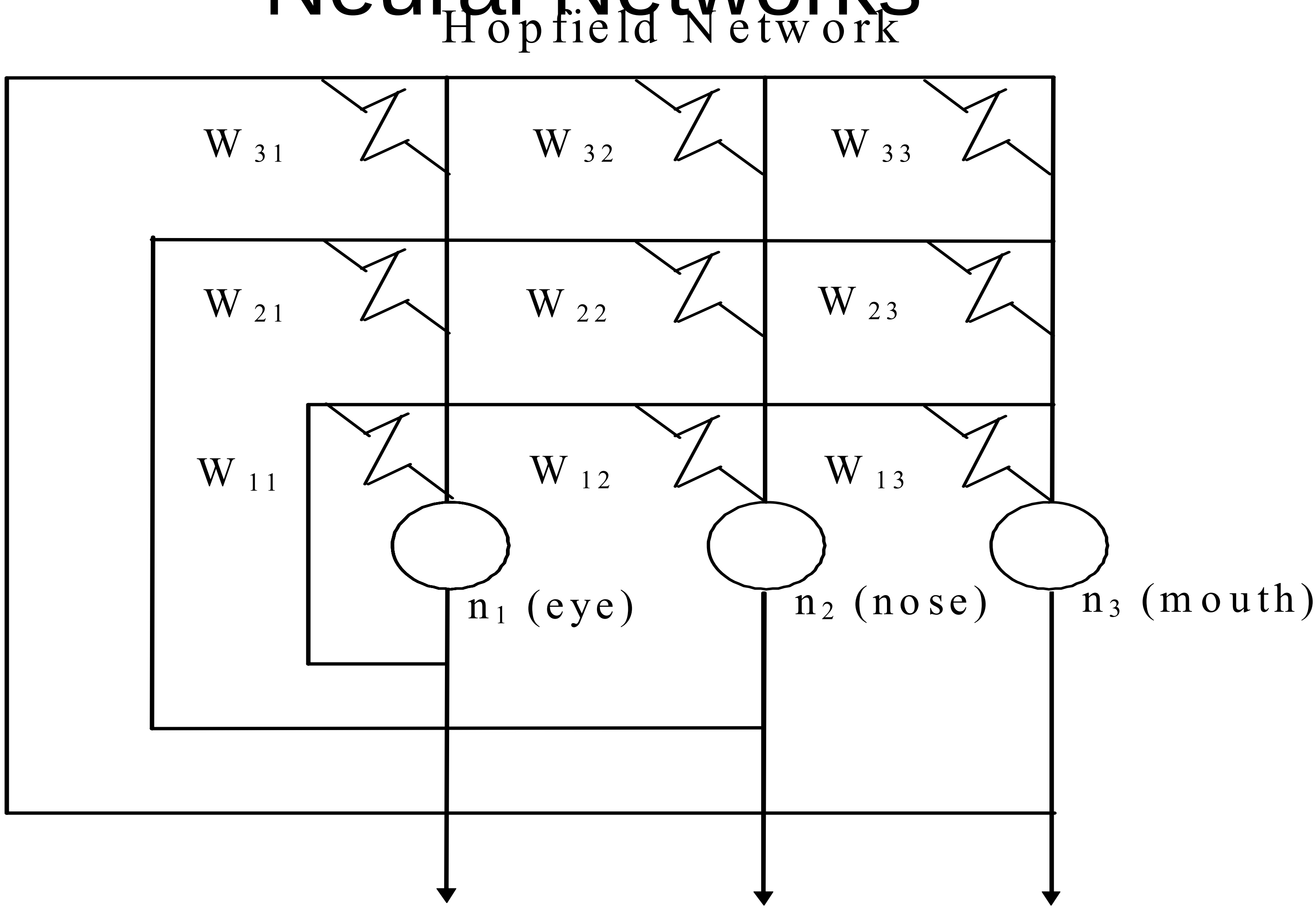
$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}^T$$



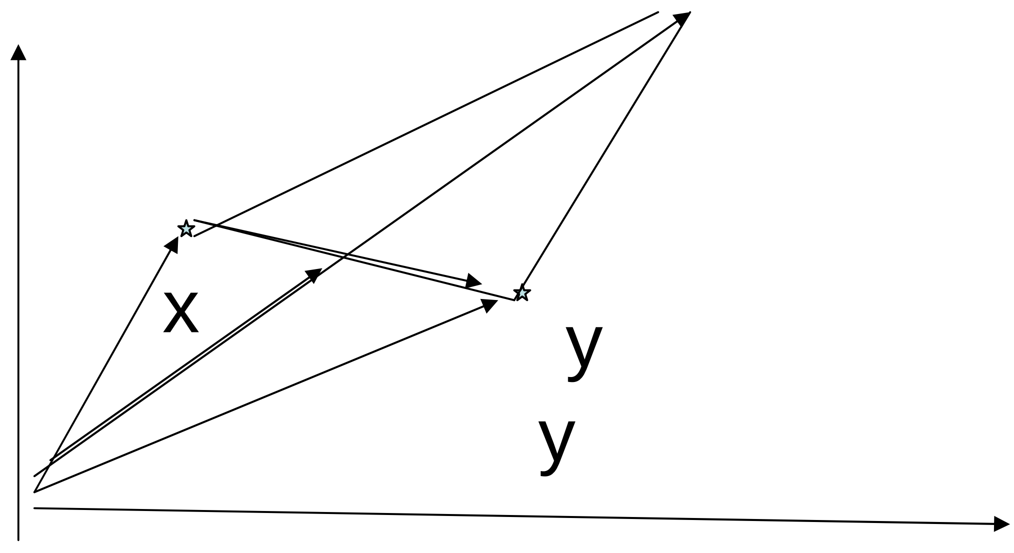
H. Szu, Brain-style Computing 6th Gen Computer Peking U. Press 1992

H. Szu and J. J. Caulfield, "**Optical Expert Systems**," *Applied Optics*, **26** (1987), pp. 1943-1947.

Fault Tolerance Search Index by Neural Networks



Self Organization Map of Kohonen



$$X_{new} = (X_{old}+Y)/2= X_{old}+ (Y- X_{old})/2$$

$$\langle X_n\rangle = \langle X_{n-1}\rangle + (1/n-1)(X_n - \langle X_{n-1}\rangle)$$

SOM is supervised i.e. a labeled data set centroid feature extraction

Kohonen Self Organization Map Centroid algorithm Batch Average versus Sequential Average

$$< X >_{m+1} \equiv \frac{1}{(m+1)w} \sum_{i=1}^{m+1} x_i w_i$$

$$= \frac{(m+1-1)}{m+1} \frac{1}{m} \sum_{i=1}^{m} x_i + \frac{1}{m+1} x_{m+1}$$

$$=< X >_m + K (x_{m+1} - < X >_m);$$

$$Kalman \quad gain \quad K = \frac{1}{m+1} :$$

$$if \quad uniform \quad weight \quad w_i = w = 1$$

1. Unsupervised Learning Dude Hart unlabelled data de-mixing find the ill-posed inverse of X=[A]S of unknown [A] and S:

2. Two breakthroughs were based on either mathematical statistics by Bell & Sejnowski of Salk, Amari of RIKEN, Oja of Finland (BSAO) since 1997 or by physics laws by Szu in 1997:

3. BSAO mathematics solution, called first by Como as **Independent Component Analysis (ICA),** assumed a single unknown impulse response function having the mixing matrix [A] for all pixels, or the space-invariant imaging (bi-annual conferences since 1998):
joint-pdf Factorization $\rho(x_1,x_2,x_3,.) = \rho 1(x'_1)\ \rho 2(x'_2)$.
Find X'=[W]X by (d [W]dt) = (d Entropy/d[W]).

4. Szu thermodynamic physics of isothermal equilibrium in the free Helmholtz energy: min. H = E(I/O) – T_o entropy S
for space-variant mixing [A] in remote & tumor using **Pixel-by-pixel Blind Source Separation (BSS)**

Historical Perspective of ANN

Auto-Regression: Given past & present data as column vector:

$X_m(t) = (x(t), x(t-1), \ldots, x(t-m+1))^T$ **U(t+1)** $= w_m(t)^T X_m(t)$

Mitigate the main limitation of Wiener AR leading to Rosenblatt & Widow ANN

(i) LMS error function supervision by a desirable output x(t+1) exemplar:

$$E = \langle (U(t+1) - x(t+1))^2 \rangle$$

(ii) Linearity, simple nonlinear saturation of the dynamic range is by using sigmoid function:

$$V(t+1) = \sigma(w_m(t)^T X_m(t)) \cong w_m(t)^T X_m(t) = U(t+1).$$

In case of imaging point spread function y(t) =s(t) * x(t) + n(t), the convolution product we relate AR to Wiener Filter.

(iii) AR is fixed-point solution at bottom of the valley:
$dw_m/dt = - dE/dw_m = 0$

We assume a piecewise, linear, and stationary *AR*.

Define stationary covariance to be difference
$C_s = \langle x(t)\, x(t-s) \rangle$

$$dE/dw_m = 2\langle [\, w_m^T X_m(t) - x(t+1)]\; X_m(t)^T \rangle = 0$$

$X(t) = (x(t), x(t-1), \ldots, x(t-m+1))^T$ *Wiener missile guidance AR 1950*

$u_i(t+1)=w_i (t)^T X(t)$ ***predicts the future from the past***

$E = < [u_i(t+1) - x(t+1)]^2>$ ***LMS***

$v_i(t+1) = \sigma(w_i (t)^T X(t)) \cong w_i (t)^T X(t) = u_i (t+1)$ ***Nonlinear ANN***

$dw_i/dt = - \partial E/\partial w_i = 2 < [w_i^T X(t) - x(t+1)] X(t)^T> =0$ ***Fixed p.***

$C_s = <x(t) x(t-s)>$ ***stationary Teoplitz matrix***

$$\begin{bmatrix} C_0 & C_1 & C_2 & \cdots & C_{m-1} \\ C_1 & C_0 & C_1 & \cdots & C_{m-2} \\ C_2 & C_1 & C_0 & \cdots & C_{m-3} \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ C_{m-1} & C_{m-2} & \cdots & C_1 & C_o \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_m \end{bmatrix} = \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ \vdots \\ C_m \end{bmatrix}$$

$dw_i/dt = -\partial E(v_i, x(t+1))/\partial w_i$ Nonstationary *ANN 1970*

$dw_i/dt= <\partial K(V)/\partial w_i >=0; K(V) = <V^4> - 3 (<V^2>)^2$

$<v_1v_2v_3v_4>_G =$ **Unsupervised Fast-ICA 2000**

$<v_1v_2>_G <v_3v_4>_G + <v_1v_3>_G <v_2v_4>_G + <v_2v_3>_G <v_1v_4>_G$

From AR Wiener to Supervised ANN, Unsupervised Fast ICA, PCA to ICA

$$\mathbf{d}\boldsymbol{w}_{\mathbf{m}}\mathbf{/dt} \;=\; \mathbf{-}\;\; \mathbf{dE(U, x(t+1))/d}\boldsymbol{w}_{\mathbf{m}}$$

$$\begin{bmatrix} C_0 & C_1 & C_2 & \cdots & C_{m-1} \\ C_1 & C_0 & C_1 & \cdots & C_{m-2} \\ C_2 & C_1 & C_0 & \cdots & C_{m-3} \\ \vdots & \ddots & \vdots & \vdots & \cdot \\ C_{m-1} & C_{m-2} & \cdots & C_1 & C_o \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_m \end{bmatrix} = \begin{bmatrix} C_1 \\ C_2 \\ .\,C_3 \\ \ddots \\ C_m \end{bmatrix}$$

Unsupervised ANN prediction by 4th order cumulant Kurtosis $\mathbf{K(V)=\langle V^4\rangle - 3\langle V^2\rangle^2}$

$$\mathrm{d}\boldsymbol{w}_{\mathrm{m}}/\mathrm{dt} \quad = \quad \mathrm{dK(V)/d}\boldsymbol{w}_{\mathrm{m}}$$

$$\mathbf{\langle X(t)X(t)^T \rangle\, e_l = \lambda\, e_l}$$

$$\mathbf{w' = w+\sigma(w^TX)X = X\sigma(X^Tw) \cong X\, X^Tw \cong .\varepsilon \langle X^TX\rangle w}$$

$$\mathbf{w'-w = \Delta w = \Delta t \langle X^TX\rangle w; \quad dw/dt = \langle X^TX\rangle w}$$

$$w' = w + \alpha\, dH(w^TX)/dw \equiv w+\alpha\, Xh(w^TX),$$

$$\mathbf{|w'|^{-1} = [(\, \boldsymbol{w+\alpha\, Xh(w^TX)^T(w+\alpha\, Xh(w^TX)}]^{-1/2} \;=\; |w|^2 \;-\; \alpha h/2(X^Tw + w^TX) + O(\alpha^2);}$$

$$\boldsymbol{Hypersphere\ \Delta w = \alpha\, [\, [I] - w\, w^T\,]\, X\, h(w^TX\,)}$$

$$\boldsymbol{Projection\ Pursue:} \quad \boldsymbol{\Delta w = \alpha\;\; h\;\; [\, [I] - \;\; w\, w^T\,]\, X \;\Rightarrow \alpha\;\; h\, [\, X - \; X\, (X^T X)] = 0}$$

Single Neuron ICA Bell & Sejnowski 1996

Unsupervised Learning by Maximizing Entropy

Boltzmann Entropy ***H*** (R, B) = Log { (R+B) !/ R! B!}

Via the Stirling Formula Log N! = NLogN – N

Is the Shannon H(R, B) = - R Log R - B Log B

$$\boldsymbol{H}(y) = -\int f(y)\,\boldsymbol{Log}\,f(y)\,dy = -\langle \mathrm{Log}\,f(y)\rangle$$

The condition norm of $1 = \int f(y)\,dy \qquad = \int g(x)\,dx$

p.d.f.: A strong condition is: $f(y) = g(x) \;/\; |dy/dx|$

$$\boldsymbol{dw/dt = \partial H/\partial w} \qquad \Delta w = \partial\boldsymbol{Log}|dy/dx|/\partial w$$

$$= dy/dx|^{-1}\,\partial\,|dy/dx|\,/\partial w$$

$$y = [\,1 + \exp(-(wx - w_o))\,]^{-1} \qquad \boldsymbol{=}\; 1/w \;-\; (2y-1)\,x$$

$$dy/dx \quad = \quad w\,y(1-y);$$

$$dy/dw \quad = \quad x\,y(1-y);$$

$$\partial|dy/dx|/\partial w \boldsymbol{=} y(1-y) + w\,\boldsymbol{(}\,dy/dw \,\boldsymbol{-}\, 2\,y\,dy/dw\,\boldsymbol{)}$$

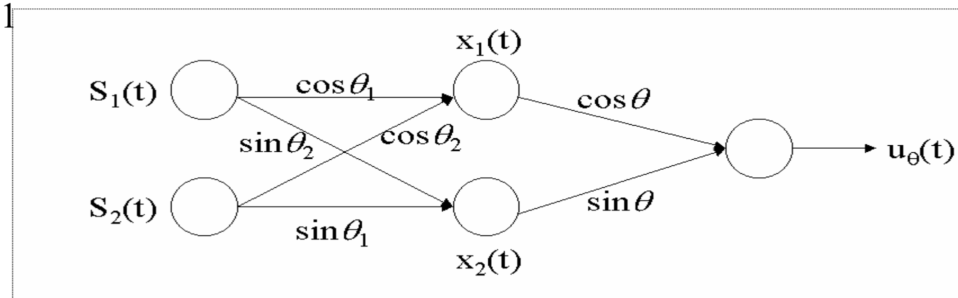
N Neurons $\boldsymbol{\partial H/\partial[w] = \Delta\,[w] \;=\; [w^T]^{-1} \;-(2y-1)\,x^T}$

$$[\boldsymbol{w}]^T[\boldsymbol{w}] = [I] = [\boldsymbol{w}]^{-1}[\,\boldsymbol{w}]$$

Amari: $\partial H/\partial[\boldsymbol{w}][I] = \{[I] - (2\boldsymbol{y}-1)\boldsymbol{x}^T[\boldsymbol{w}]^T\}\,[\boldsymbol{w}] = \{[I] - (2\boldsymbol{y}-1)\boldsymbol{u}^T\}[\boldsymbol{w}]$

Example $\vec{x}(t) = \sum_{i=1}^{2} \vec{a}_i s_i(t) = [A]\vec{s}(t)$

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} \cos(\theta_1) & \cos(\theta_2) \\ \sin(\theta_1) & \sin(\theta_2) \end{pmatrix} \begin{pmatrix} s_1(t) \\ s_2(t) \end{pmatrix}$$



$$u_\theta(t) = [\cos\theta \quad \sin\theta] \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \vec{w}^T \vec{x}(t)$$

$$= \sum_{i=1}^{2} (\cos\theta \cos\theta_i + \sin\theta \sin\theta_i) s_i(t)$$

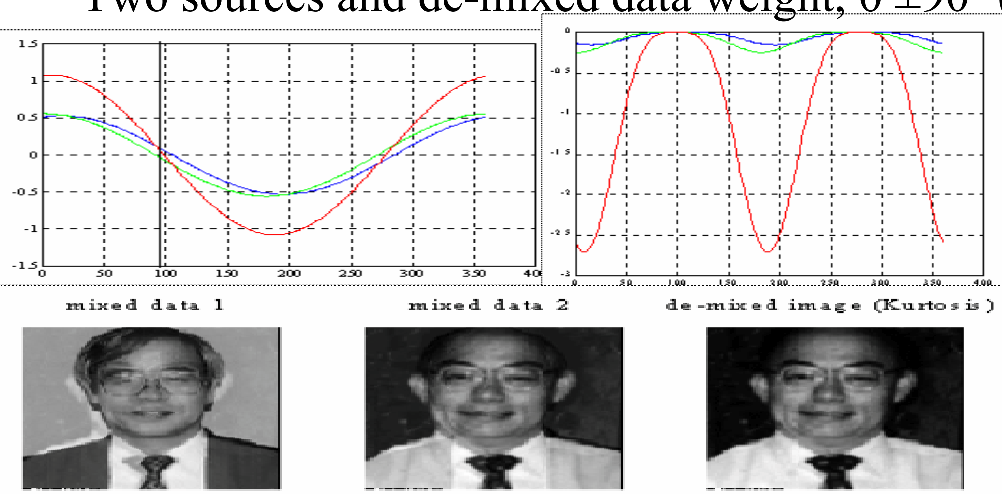
$$= \sum_{i=1}^{2} \cos(\theta - \theta_i) s_i(t)$$

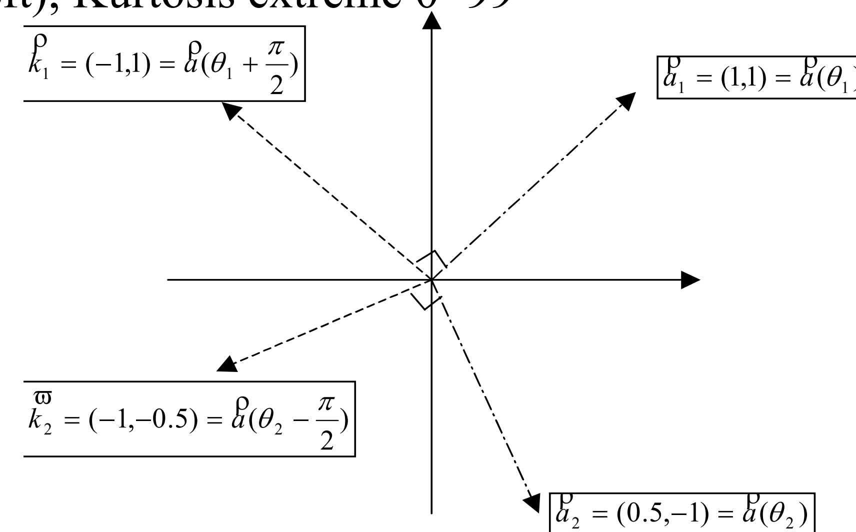
$$\begin{bmatrix} u_{\theta 1} \\ u_{\theta 2} \end{bmatrix} = \begin{bmatrix} 1 & \cos(\theta_1 - \theta_2) \\ \cos(\theta_2 - \theta_1) & 1 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}$$

$$K(u_{\theta 1}) = K(s_1) + \cos(\theta_1 - \theta_2) K(s_2)$$

Unknown θ_i (blind) if it happens at the killing angles, $\theta = \theta_1 + \pi/2$ then it gives $\cos(\pi/2)=0$ and $u_\theta(t)=\sin(\theta_2 - \theta_1)s_2(t)$.

Two sources and de-mixed data weight, θ ±90° (left); Kurtosis extreme θ=99°





Comparison: Single pixel E([W]X-S) in isothermal equilibrium min. $H=E-T_oS$, versus ICA Max. Post Entropy for a fixed [A] for all pixel ensemble averaging

•Space-variant imaging (Variable response [A] pixel to pixel)

•Pixel-parallel independent search

Szu applied min. Helmholtz free energy per pixel to reduce the uncertainty among many inverse solutions.

Min. Helmholtz $H=E([W]X-S) - T_oS(s)$

Shannon-Boltzmann a priori Entropy:

$$S(s_1,s_2,s_3,.)= -\Sigma s_i \log s_i+(\mu_o+1)(\Sigma s_i-1)$$

by postulating the first order estimation error information I/O energy

$$E=\mu([W]X-S)=\lambda([A]S-X)$$

derived ANN sigmoid and Hebb rule:

Derived: 1. Sigmoid, 2. Unsupervised learning Hebb rule 3. Real world applications Remote sensing, breast cancer

•Space invariant imaging (Identical response [A] of a closed system)

•advantage of pixel ensemble

• Bell-Sejnowski, Amari, Oja (BSAO) find ICA defined by joint pdf factorization

$$\rho(x_1,x_2,x_3,.) \Rightarrow \rho_1(x'_1)\ \rho_2(x'_2)\ldots$$

$$\frac{\partial [W]}{\partial t} = \left\langle \frac{\partial PostEnt(\vec{y}([W]\vec{x}))}{\partial [W]} [W]^T [W] \right\rangle_{pixels}$$

$$d_{\vec{y}} = (\vec{y}, \vec{y}) = (\vec{x}, [W]^T [W] \vec{x}) = d_{\vec{x}}$$

BSAO assume ANN post-processing closed system

Max Post-Entropy($V = \sigma([W]X)$)

Challenges remain:

1. Component Permutation?

2. Inhomogeneous pixel *[A]*?

3. Nonlinear ICA?

4. Biological meaning of binding?

References: Szu et al., SPIE Wavelets 1997; Szu et al.,WCCI Hawaii 2002; IEEE Trans Geoscience Remote Sensing 2002

No difference Space Invariant Imaging, single [A]

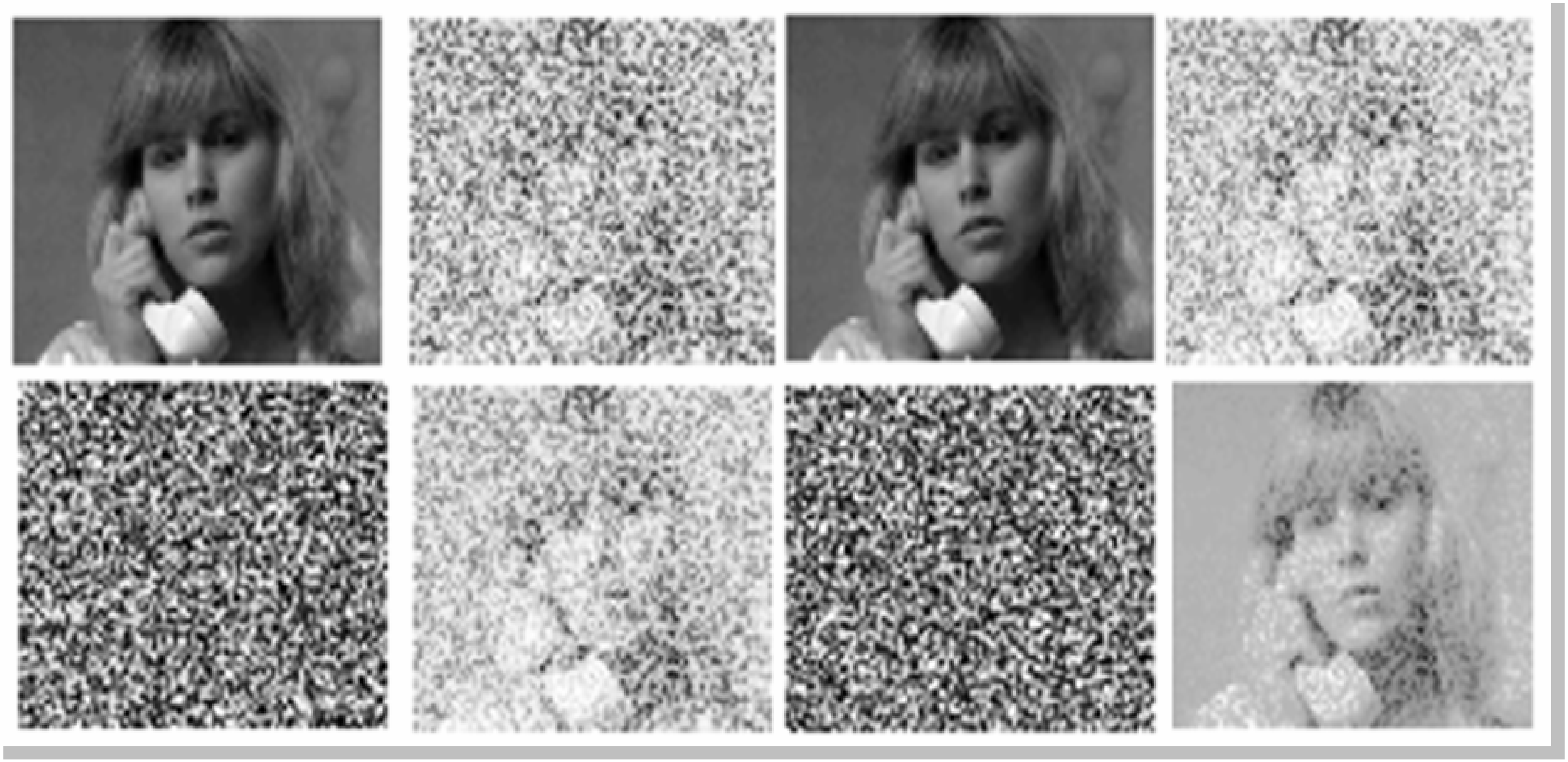
From left to right: (i) source images;
(ii) linear space-invariant mixtures;
(iii) recovered images using **linear LCNN** algorithm
(iv) recovered images using BSAO **Infomax ICA** algorithm

Space Variant Imaging, variant [A]

From left to right: (i) source images;
(ii) linear space-variant mixtures;
(iii) recovered images using **LCNN** algorithm
(iv) recovered images using **Infomax ICA** algorithm

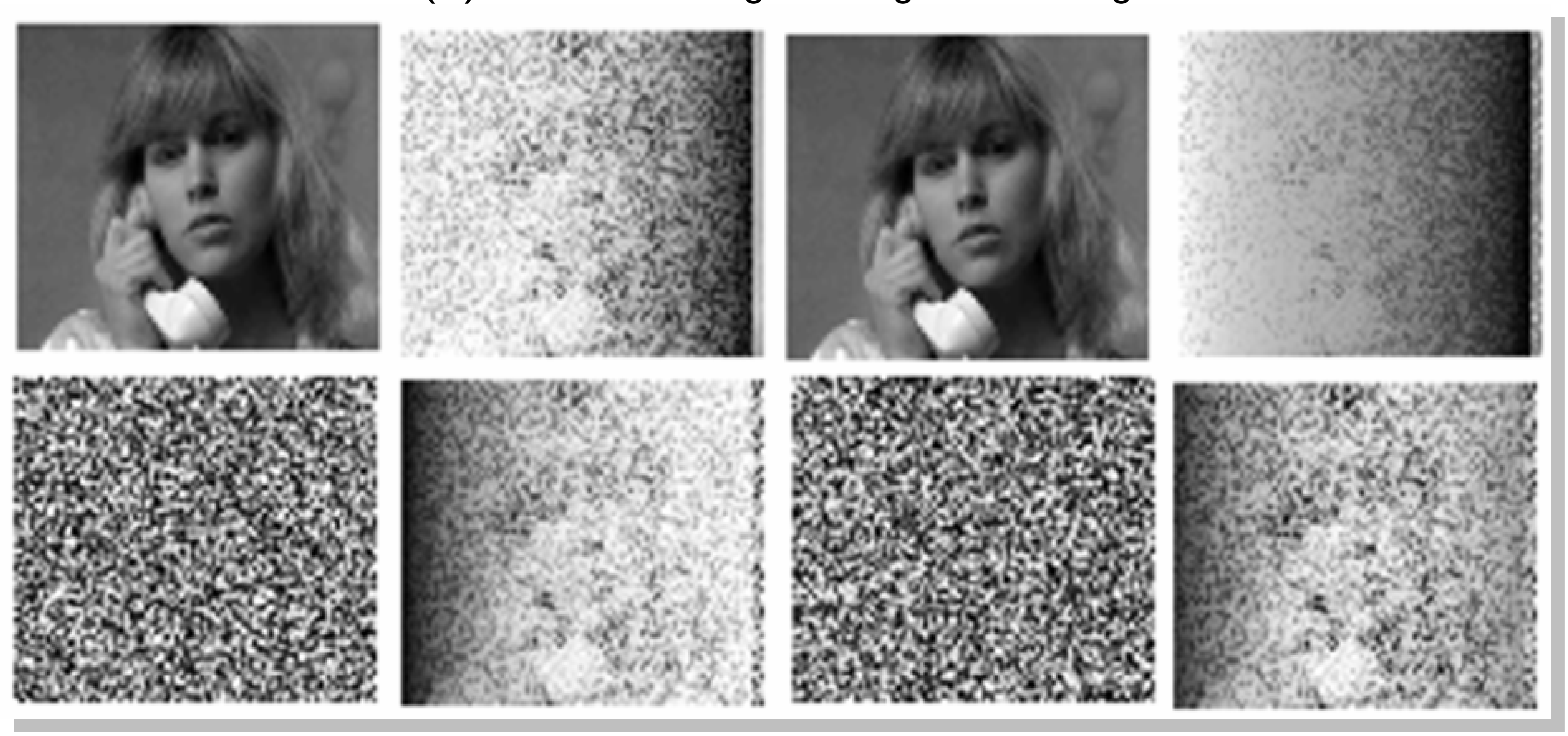
Nonlinear Space Invariant Imaging, single[A]

From left to right: (i) source images;
(ii) nonlinear space-invariant mixtures;
(iii) recovered images using **nonlinear LCNN** algorithm
(iv) recovered images using **Infomax ICA** algorithm



Nonlinear Space-variant Imaging, variant [A]

From left to right: (i) source images;
(ii) nonlinear space-variant mixtures;
(iii) recovered images using nonlinear LCNN algorithm
(iv) recovered images using Infomax algorithm

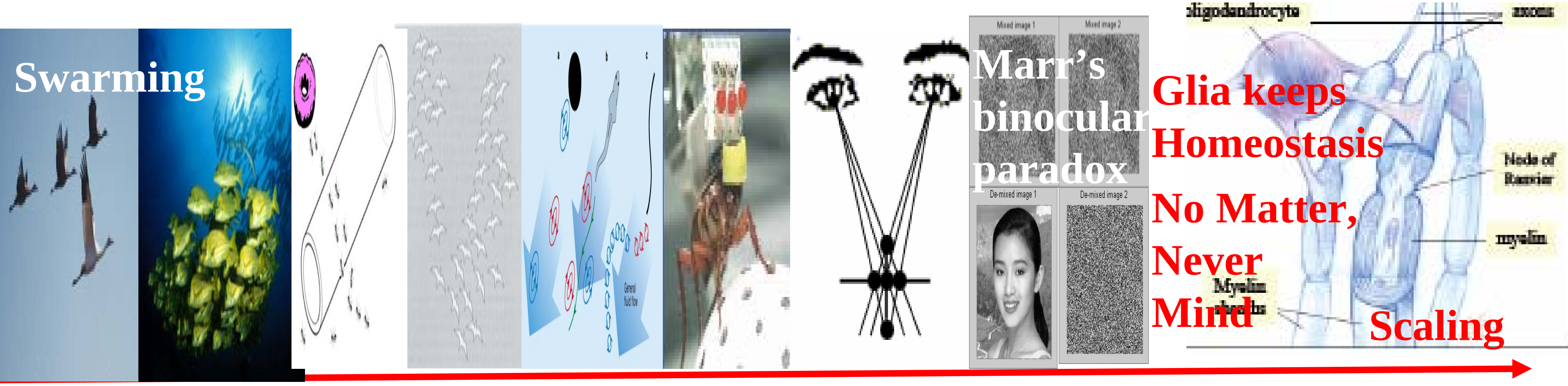


Claustrum is the legacy Crick left behind:

"No Matter, Never Mind," Is Claustrum the matter in mind?

Crick & Koch identified a thin sheet of gray matter resides in two way parallel connection below the cortex, computing feelings, seeing, hearing, language, deciding what to do. They said: ***"Claustrum is a conductor of consciousness like an orchestra that binds all senses". Philos. Trans. Royal Soc. Lond, 2005***

From Consciousness to mind from mind to body vice versa in society



From Physics To Physiology estimates 20 B of Neurons

H. Szu 2007

•If one takes Crick-Koch conscious model---***Claustrum*** seriously, then the ubiquitous binding of bursting 30Hz firing rates is the wakening of consciousness.

•What would be the size of neurons & house-keeping glia cells in our brain kept at 37°C that can support effortlessly? It turns out to be tenths billions.

•Owing to Boltzmann and Einstein formulae, our equilibrium brain size may be estimated as follows: ((1/40) $eV = K_B T_o$ at 300°K room temperature);

•*E= Nhf =* $K_B T_o$*,ranging hundredth to tenth* ***billions glias & neurons*** *corresponding to* ***30Hz and 60 Hz*** *firing rates respectively for the ground state of mindlessness.*

$$
\begin{aligned}
E &= N\,hf = N\,6.67 \times 10^{-34} f \;\; Joule \\
&\approx K_B T_o = 1.38 \times 10^{-23} \times 310 \;\; Joule \\
Thus,\; N &\le 2 \times 10^{11}, \;\; if \;\; f = 30 \;\; Hz \\
N &\le 10^{11}, \;\; if \;\; f = 60 \;\; Hz \\
N &\cong 6 \times 10^{10}, \;\; if \;\; f = 100 \;\; HZ
\end{aligned}
$$

•One shall tabulate the sizes of brains with respect to the temperatures kept by **warm blood animal kingdom** for neuroethological and development mechanism.

•Far from equilibrium life forms and evolves. Math-Physics waits for new method.

•Equilibrium physics value provides local insight and boundary condition.

Long Term Neurointelligence: physiology & physics elucidate, exploit sub-consciousness (sublimated awareness) **to early perception** (reasoning, memory) **self learning** (supervised, **unsupervised**, decision, action) ***swarm robotic intelligence*.**

Mid Term: **Substantiate** house-keeping ***glia cells "μ_j"functionality*** as the ***missing half of Einstein's brain***, maintaining synaptic junctions providing. "(1) Why do we have pairs of sensors? (2) Why constant temperature brains? What is the function and behavior of glia at individual synapses? Should the interactions among glia cells and neurons follow the Gibbs spontaneity:min. $H = E_0 + \sum_i \mu_i \{[W_{ij}]X_j(t) - s_i(t)\} - T_o S$ where $T_o = 37^0C$; S= entropy; E=I/O energy? We derived well known

(i) Hebb Rule: $$\frac{\partial W_{ij}}{\partial t} = -\frac{\partial H}{\partial W_{ij}} \Rightarrow \Delta W_{ij} \approx \mu_i X_j$$

(ii) Sigmoid Rule: $$\frac{\partial H}{\partial s_i} = 0 \Rightarrow \sigma(\mu_i) = s_i$$

Near Term: To build test-bed of **biomimetic pairs of biosensors** and to demonstrate a homeostasis **Unsupervised Learning** capability for signal and noise enhancement, & feature extraction for **Biomedical Wellness (BMW)** engineering.

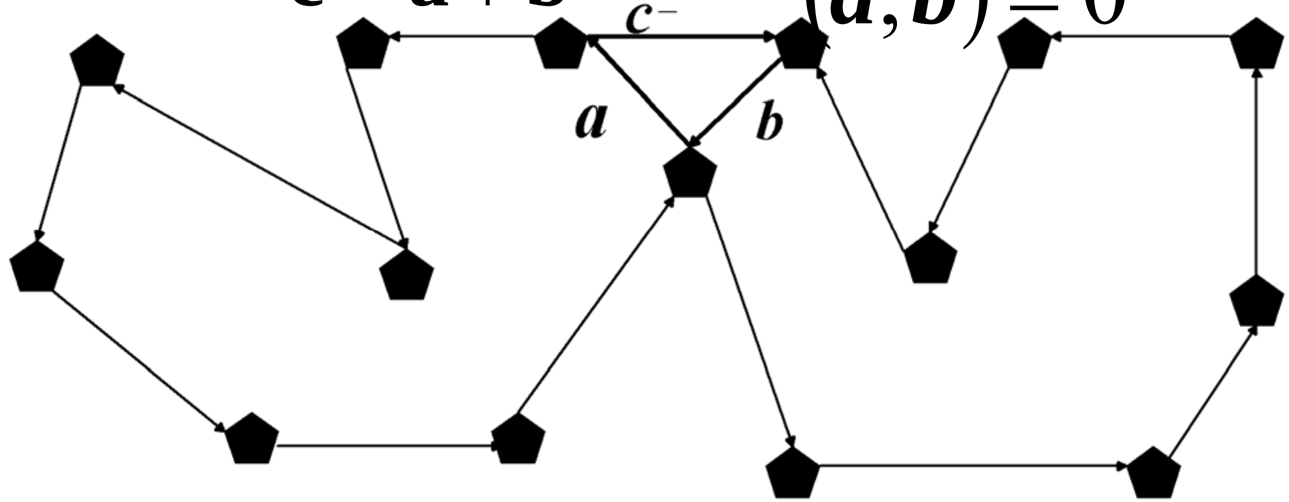
High-Impact: **Interdisciplinary Computing Intelligence (ICI)** serve the wellness of aging population measuring daily the "Wellness Baseline Profiling (WBP)" to save healthcare cost by improving the quality of life of ¼ US Population (78M aging baby boomers) costing ¼ GDP with the earliest possible diseases and dementia.

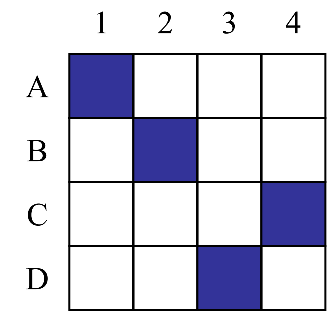
Classical ANN Constraint Optimization by Divide & Conquer Solving N-P Complete TSP

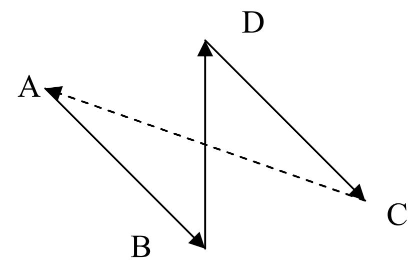
A vector tour sum = minus-one city tour was decomposed into two $\boldsymbol{a}, \boldsymbol{b}$ separate minus-one city tours such that an orthogonal decomposition could be found.

$$D = \sum_{j=1}^{N} \sqrt{(\boldsymbol{c}_j, \boldsymbol{c}_j)} \qquad \sum_{j=1}^{N} \boldsymbol{c}_j = 0 \qquad \mathbf{c} = \sum_{j=1}^{N-1} \mathbf{c}_j = -\,\mathbf{c}_N$$

$$\mathbf{c} = \mathbf{a} + \mathbf{b} \qquad (\boldsymbol{a}, \boldsymbol{b}) = 0$$







Theorem $\min(\mathrm{c}, \mathrm{c}) = \min(\mathrm{a}, \mathrm{a}) + \min(\mathrm{b}, \mathrm{b}) + 2\,(\mathrm{a}, \mathrm{b}) = \min(\mathrm{a}, \mathrm{a}) + \min(\mathrm{b}, \mathrm{b})$

$$E = \frac{1}{2}\sum_i\sum_j\sum_k d_{i,j}(S_{i,k}S_{j,k+1} + S_{i,k}S_{j,k-1}) + c\sum_i\Big(\sum_j S_{i,j} - 1\Big) + c''\sum_j\Big(\sum_i S_{i,j} - 1\Big) + c'''\sum_i\sum_j S_{i,j}(S_{i,j} - 1)$$

Szu, International Joint Conf. on Neural Networks, Washington, DC, pp. I-507-511, June 1989.

Q.E.D.

Theorem Cauchy vs. Gaussian cooling schedules

$$T_G(t) = T_0 / \log t \qquad T_C(t) = T_0 / t$$

Proof. By negate the converse strategy

Let the state-generating probability at the time t bounded below be g_t

Then the probability of not generating by $0 \leq (1 - g_t)$

To prove a cooling schedule is to prove the impossibility of never generating a state vanished: equivalent to prove:

$$\prod_{t=t_0}^{\infty}(1 - g_t) = 0 \quad (\log 0 = -\infty, \log(1 - g_t) \approx -g_t) \quad \sum_{t-t_0}^{\infty} g_t = \infty$$

$$T_G(t) = T_0 / \log t$$

$$g_t \approx \exp\left[-|\Delta x_0|^2 / T_G(t)\right] T_G(t)^{-D/2} \qquad \sum_{t=t_0}^{\infty} g_t \geq \exp(-\log t) = \sum_{t=t_0}^{\infty} 1/t = \infty$$

$$T_C(t) = T_0 / t$$

$$g_t \approx \frac{T_C(t)}{\left[T_C^2(t) + |\Delta x_0|^2\right]^{(D+1)/2}} \approx \frac{T_0}{t|\Delta x_0|^{D+1}} \qquad \sum_{t=t_0}^{\infty} g_t \approx \frac{T_0}{|\Delta x_0|^{D+1}} \sum_{t=t_0}^{\infty} \frac{1}{t} = \infty$$

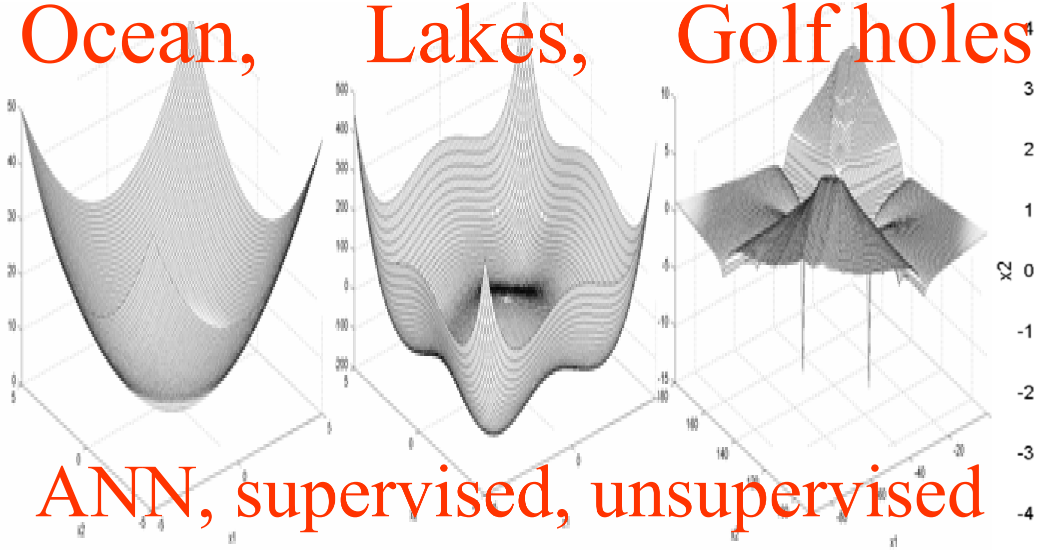
So neighborhood was visited infinite number of times at each time t for admissible cooling schedule.

Geman & Geman PAMI-6, pp. 721-741, Nov. 1984.
SZU & Hartley, Phys Lett. *A* 122, pp.157-162, 1987

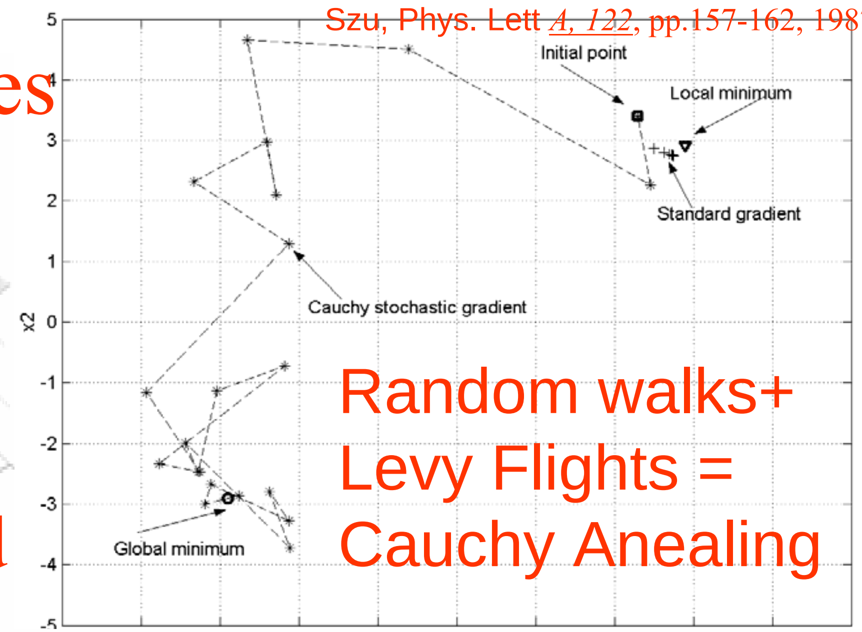
Learning Landscapes & Cauchy Annealing

Szu, Phys. Lett *A, 122*, pp.157-162, 1987

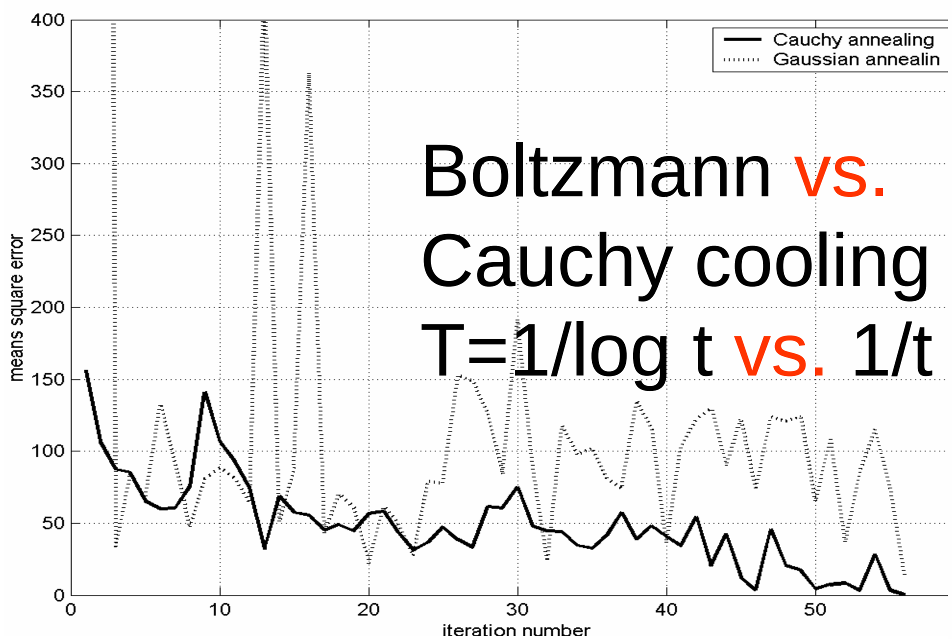
Ocean, Lakes, Golf holes



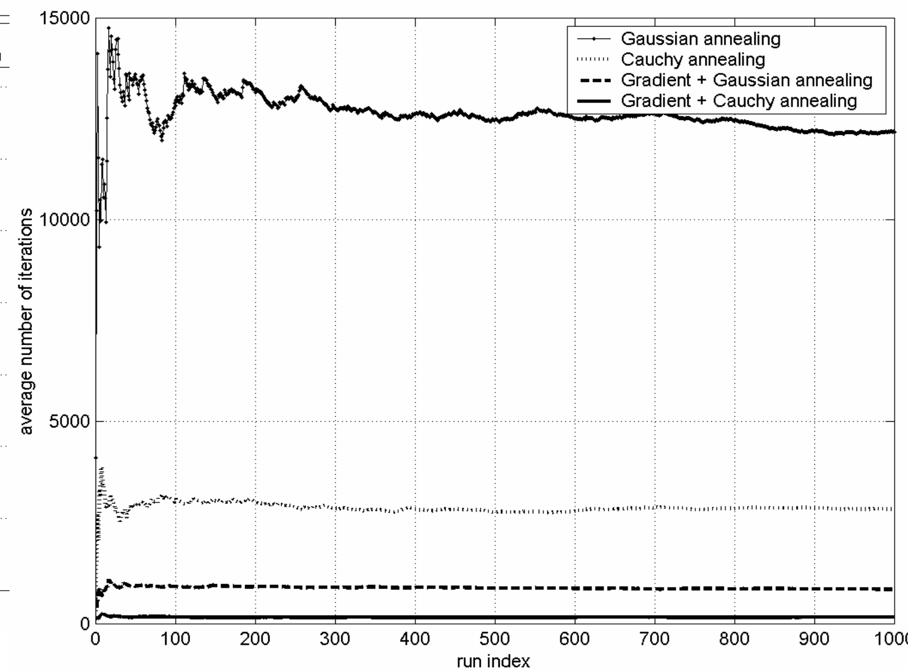
ANN, supervised, unsupervised



Random walks+
Levy Flights =
Cauchy Anealing



Boltzmann vs.
Cauchy cooling
T=1/log t vs. 1/t



Review of Stochastic Newtonian ANN Models & Convergence Proof

$$\frac{du_i}{dt} = F_i(t) \equiv < F_i(t) > + \{F_i(t) - < F_i(t) > \} \; ; \quad \frac{du_i}{dt} + \eta u_i = \widetilde{F}_i(t);$$

Langevin

$$\frac{du_i}{dt} + \eta u_i = -\frac{\partial E}{\partial v_i} + \widetilde{F}(t); \; v_i = \sigma(u_i); \; u_i = \sum_j W_{ij} v_j - \theta_i \quad W_{ij} \propto v_i u_j$$

$$\frac{du'_i}{dt} = -\frac{\partial E'}{\partial v'_i} + \widetilde{F}'(t); \; v'_i = \sigma(u'_i); \; u'_i = \sum_j W'_{ij} v'_j - \theta'_i; W'_{ij} \propto v'_i u'_j$$

$$u_i' \equiv u_i \exp(\eta t); \; v_i' \equiv v_i \exp(\eta t); \; E' \equiv \exp(2\eta t) E; \; \widetilde{F}' = \exp(\eta t) \widetilde{F} :$$

Input Hopfield dynamics

$$u_i - \sum_j W_{ij} v_j = 0; \; \Rightarrow du_i \big/ dt = -\alpha(u_i - \sum_j W_{ij} v_j); \quad \Rightarrow du_i \big/ dt + \alpha u_i = -\frac{\partial E}{\partial v_i}$$

Output Grossberg dynamics

$$v_i - \sigma(u_i) = 0; \quad dv_i \big/ dt = -\beta(v_i - \sigma(u_i))$$

Hebbian ICA dynamics a la Amari-Cichocki-Yang metric

$$\frac{dW_{ij}}{dt} = -\frac{\partial \, supervised \;\; E}{\partial W_{\alpha\beta}} g_{ij}^{\alpha\beta}; \frac{dW_{ij}}{dt} = \frac{\partial \, unsupervised \;\; S/K}{\partial W_{\alpha\beta}} g_{ij}^{\alpha\beta} = 0 \Rightarrow f.p. Fast\,ICA$$

Convergence Theorem proved via Lyaponov style: (real#)2=positive

$$\frac{dE'}{dt} = \sum_j \frac{\partial E'}{\partial v'_j} \frac{dv'_j}{du'_j} \frac{\partial u'_j}{\partial t} + \sum_i \sum_j \frac{\partial E}{\partial W_{ij}} \frac{dW_{ij}}{dt}$$

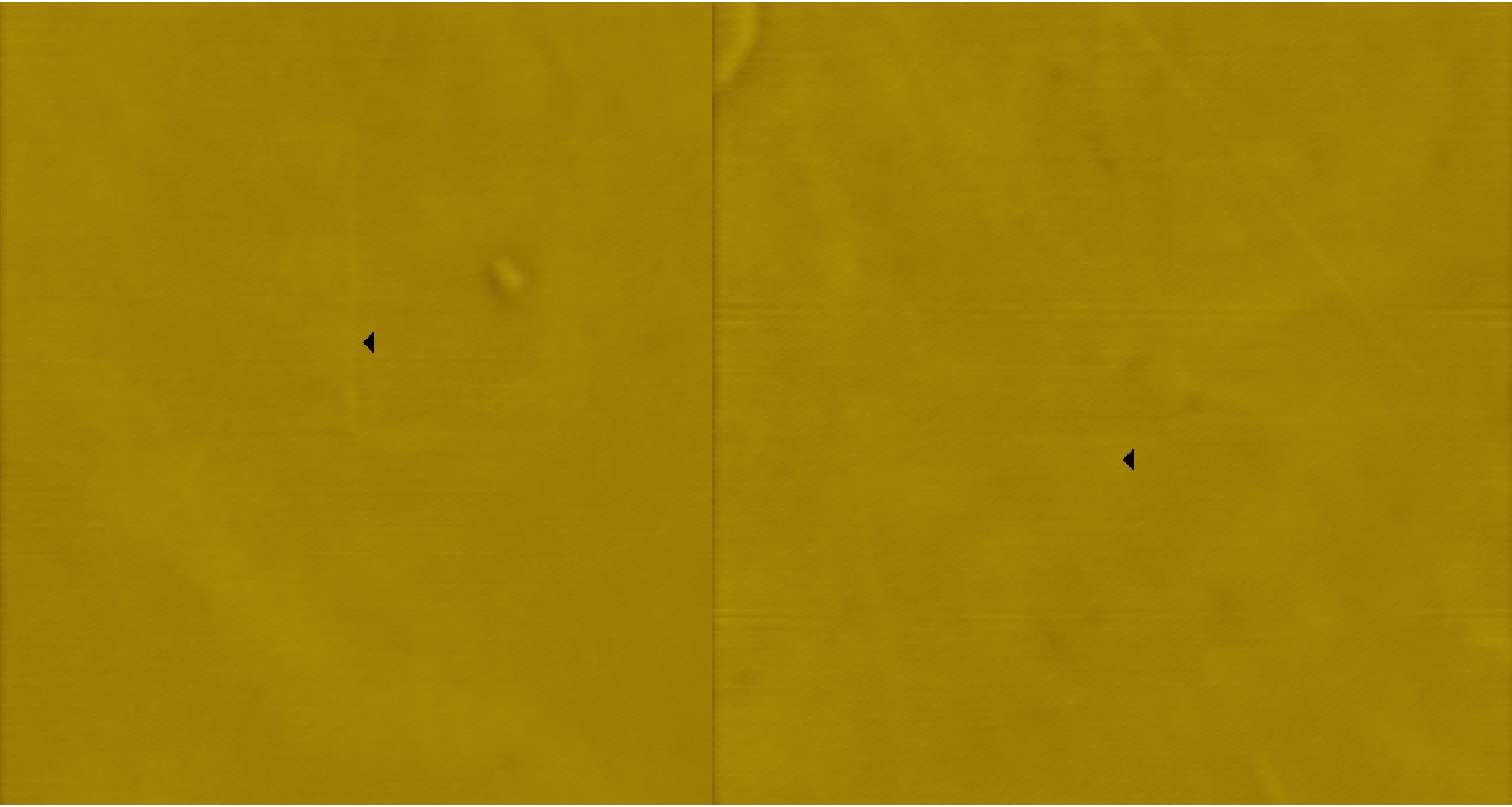
$$= -\sum_j \left(\frac{\partial E'}{\partial v'_j} \right)^2 \frac{dv'_j}{du'_j} + \left(\frac{\partial E'}{\partial v'_j} \right) \frac{dv'_j}{du'_j} \widetilde{F}'(t) - \sum_i \sum_j \left(\frac{\partial E}{\partial W_{ij}} \right) \frac{\partial E}{\partial W_{\alpha\beta}} g_{ij}^{\alpha\beta} \leq 0$$

Szu 1999 Q.E.D.

2. Real World Applications from Tank to Tumor need unsupervised MIQ/CI capability

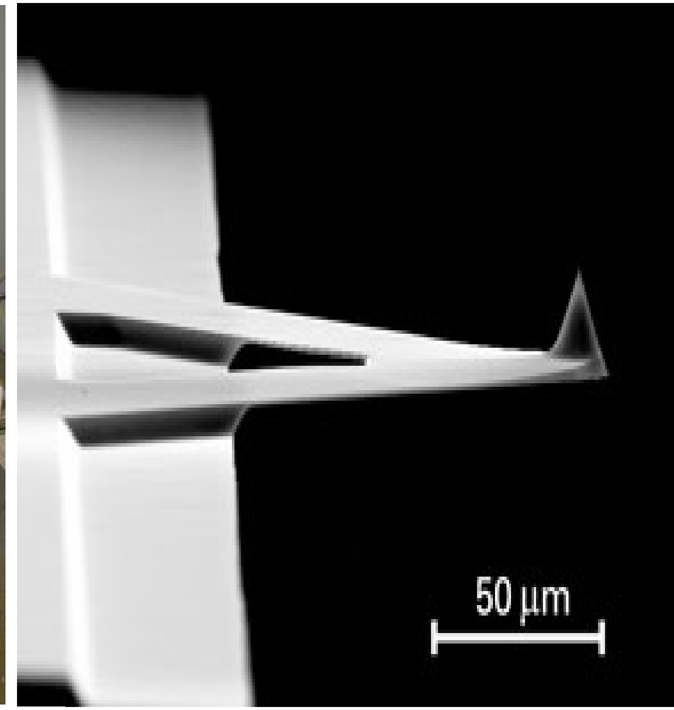
- NASA remote sensing in outer space without ground truth
- Mission To Moon 2010, To Mars !
- mini-UAV for Future Naval Capability e.g. Protect $2B Aircraft from $1M Cruise Missiles attack
- mini-UAV 3D Synthetic Aperture Radar seeing through forest, rainy foggy weather, sand storm
- NIH/NCI; NIH/NIBIB; molecular tagged imaging, early tumor detection

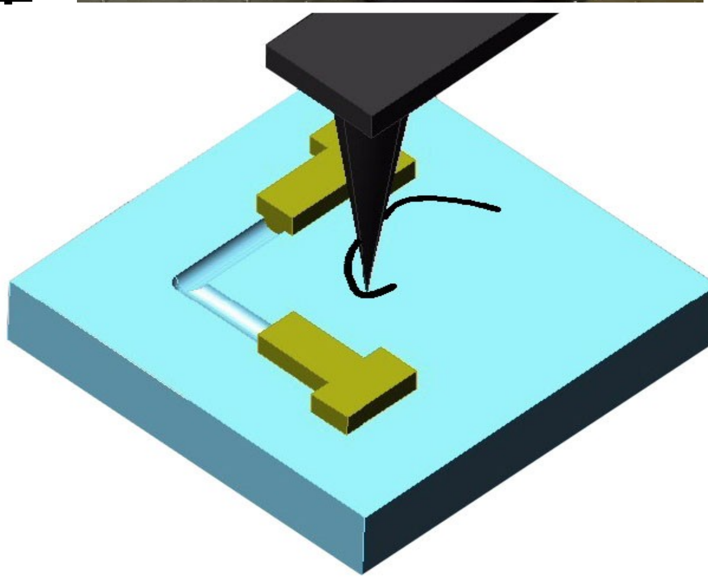
NanoRobot For Mid IR Camera

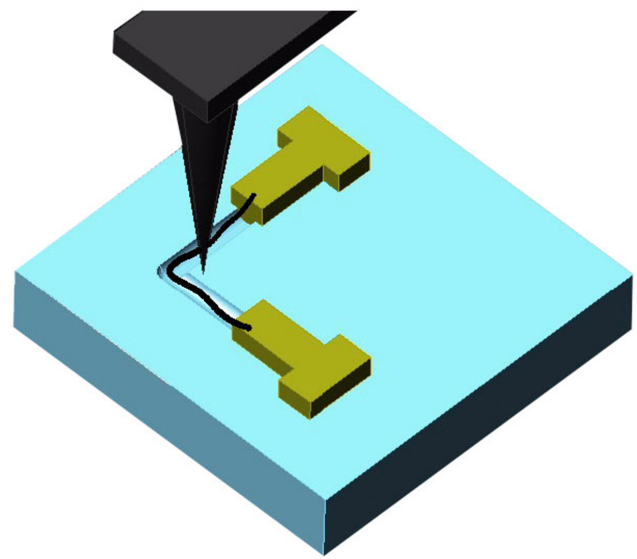


Fabrication of Sensor Pixels

- AFM based nanoassembl y
- Pushing Nanotube into nanofixture
- Tuning CN band gap b modifying i shape

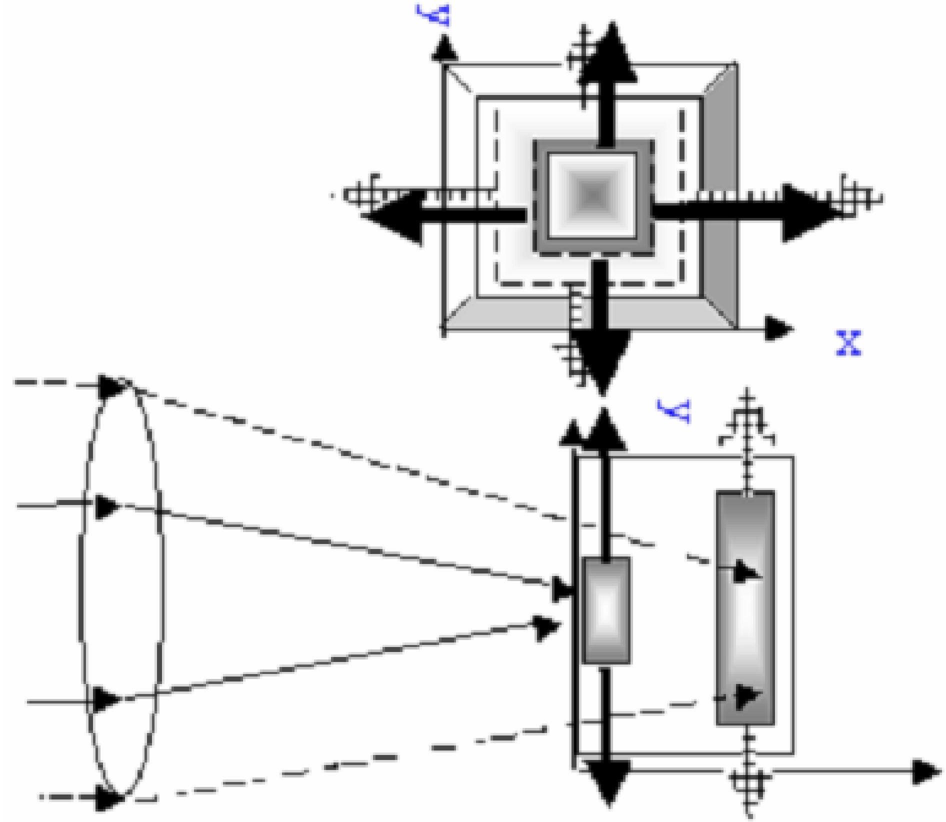






Non-Cryogenic Co-Axial Fovea Design of Infrared Two Color Pixels Planes*

- full-band IR lens focused at dual focal planes along uniaxial for both MidIR &LongIR
- Band-selective Carbon NanoTube (CNT) generates only 1% or less occlusion over Long IR PFA CCD pixel.
- CNT might only suffer 1D noise that permits a non-cryogenic cooling.
- Un-cool FPA CCD for LongIR & steering & pointing at ROI
- Unsupervised fusion for unbiased feature extraction.



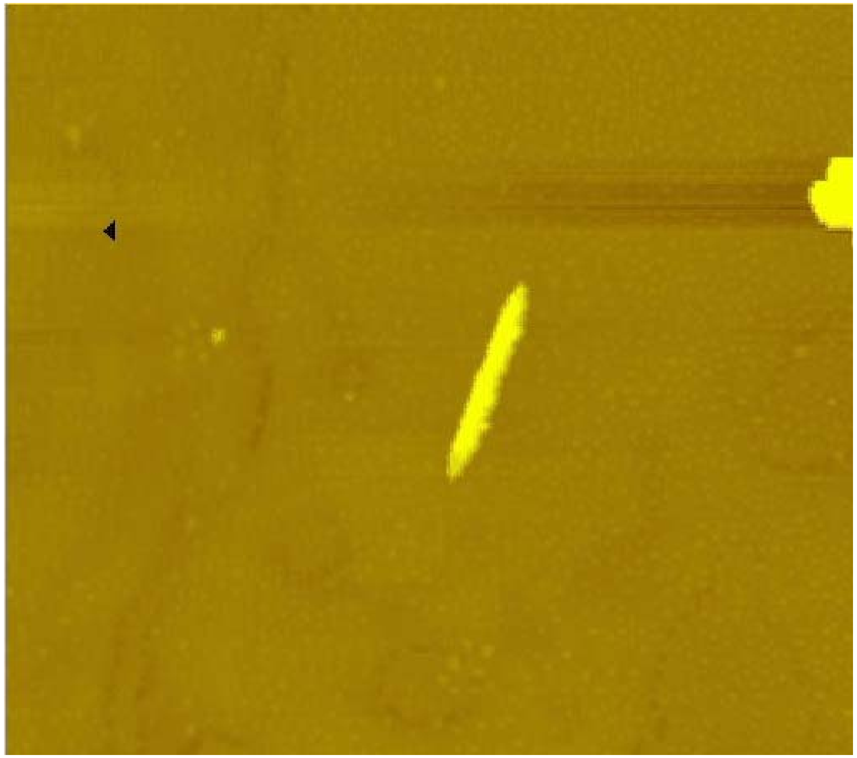
Uncooled
night vision

CCD @$1K

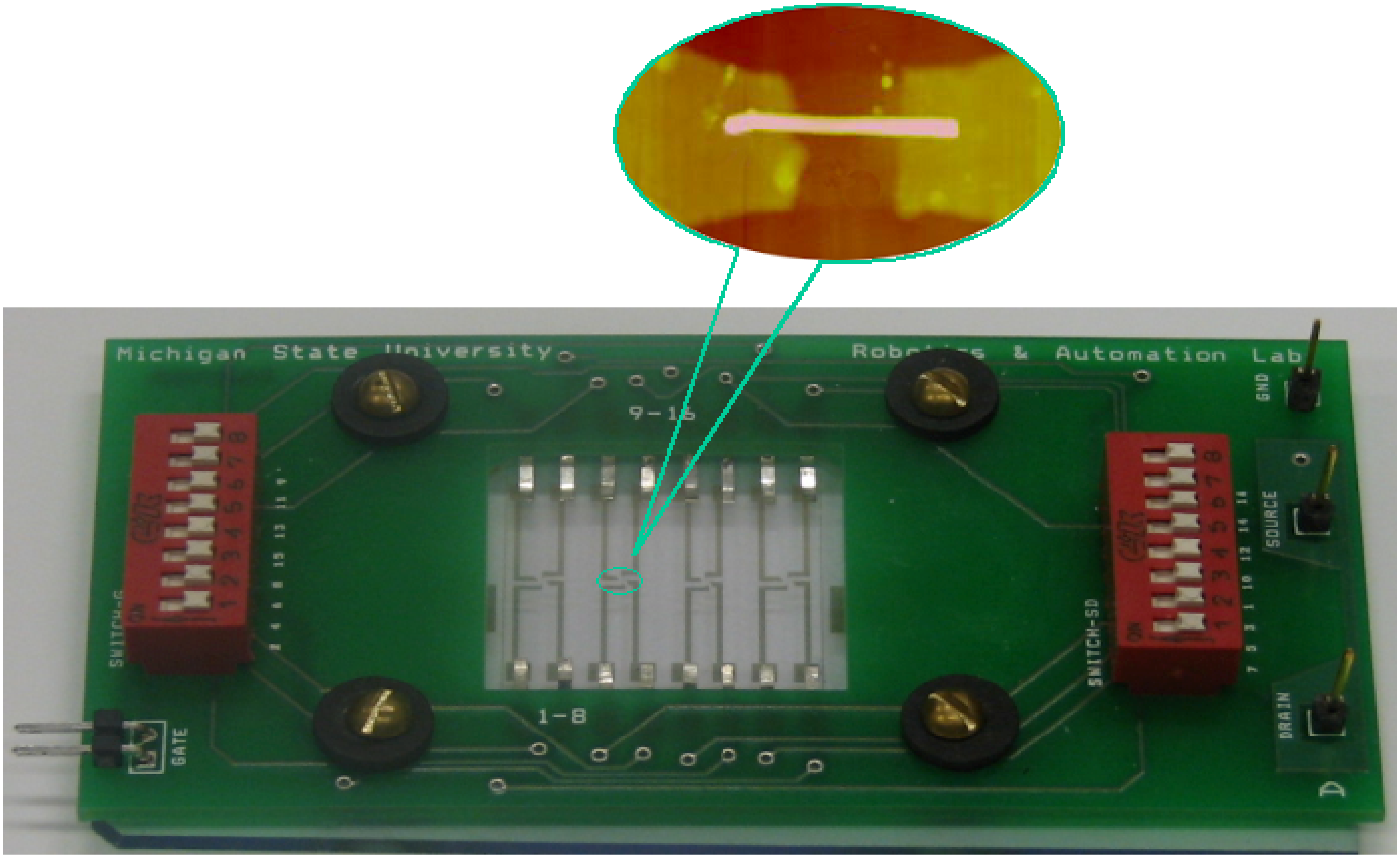
*Szu et al. Patent Disclosure 2004

NanoRobot assembly Experiment

- Creating a fixture and then pushing a 100nm silver nanowire into the fixture (scanning size 5um)

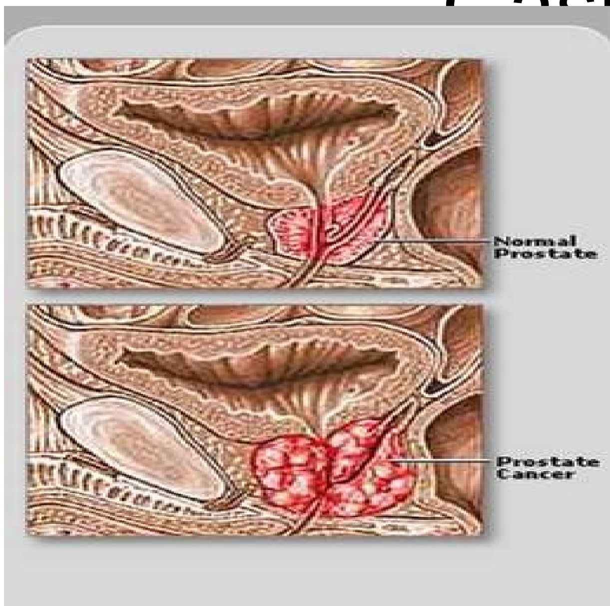


AFM image   Real-time visual display

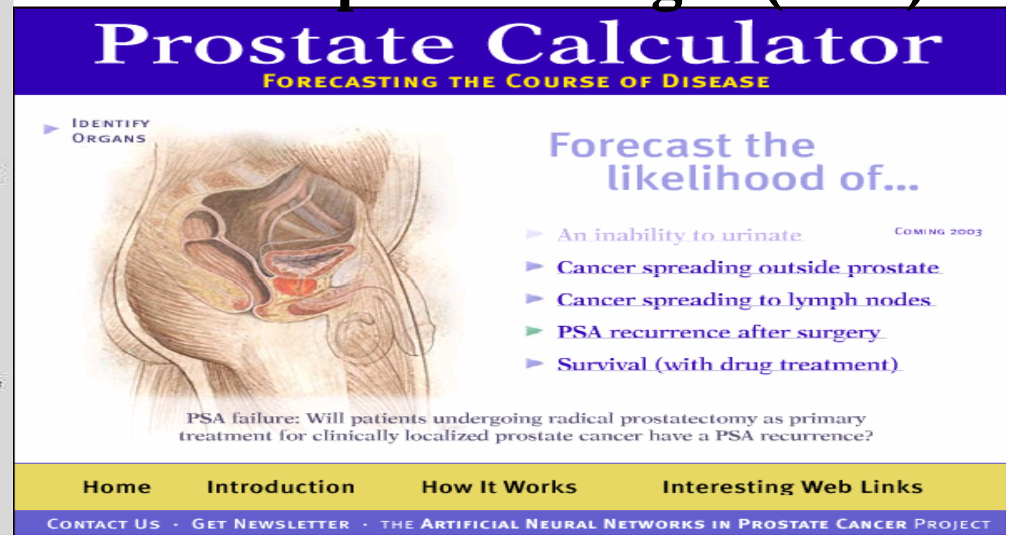


The picture shows a carbon nano tube (CNT) based IR detector array. Each pixel of the IR detector array consists of a multi-wall CNT with a proper tuned bandgap for detection of selected spectrum of the infra red.

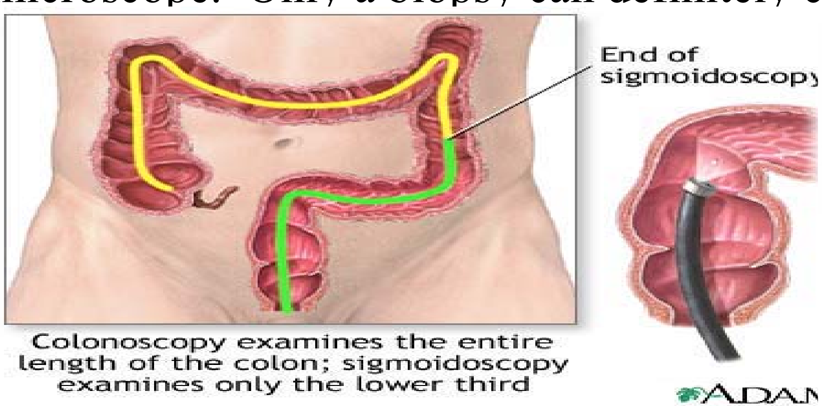
Passive IR Flexi-Scope with Two Spectral Colors for Household Screening of Gastrointestinal Disorders



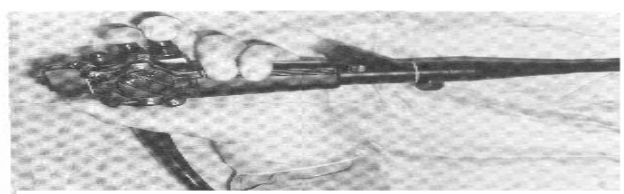
Prostate Specific Antigen (PSA)

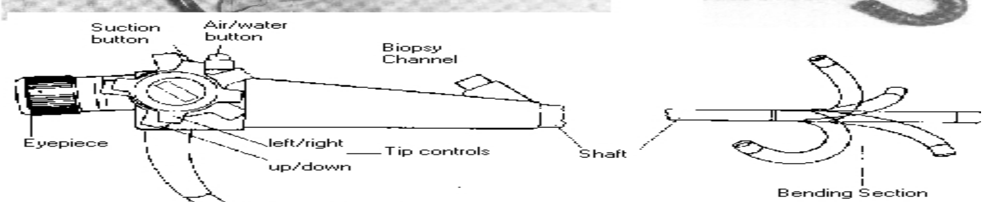


Biopsy, under anesthesia, takes a set of 2 sample places per quadrants of 8 quadrants of the prostate; a total ranges anywhere from 16 to 20 sampling for culture growth pattern under the microscope. Only a biopsy can definitely confirm the presence of prostate cancer, not vice versa.

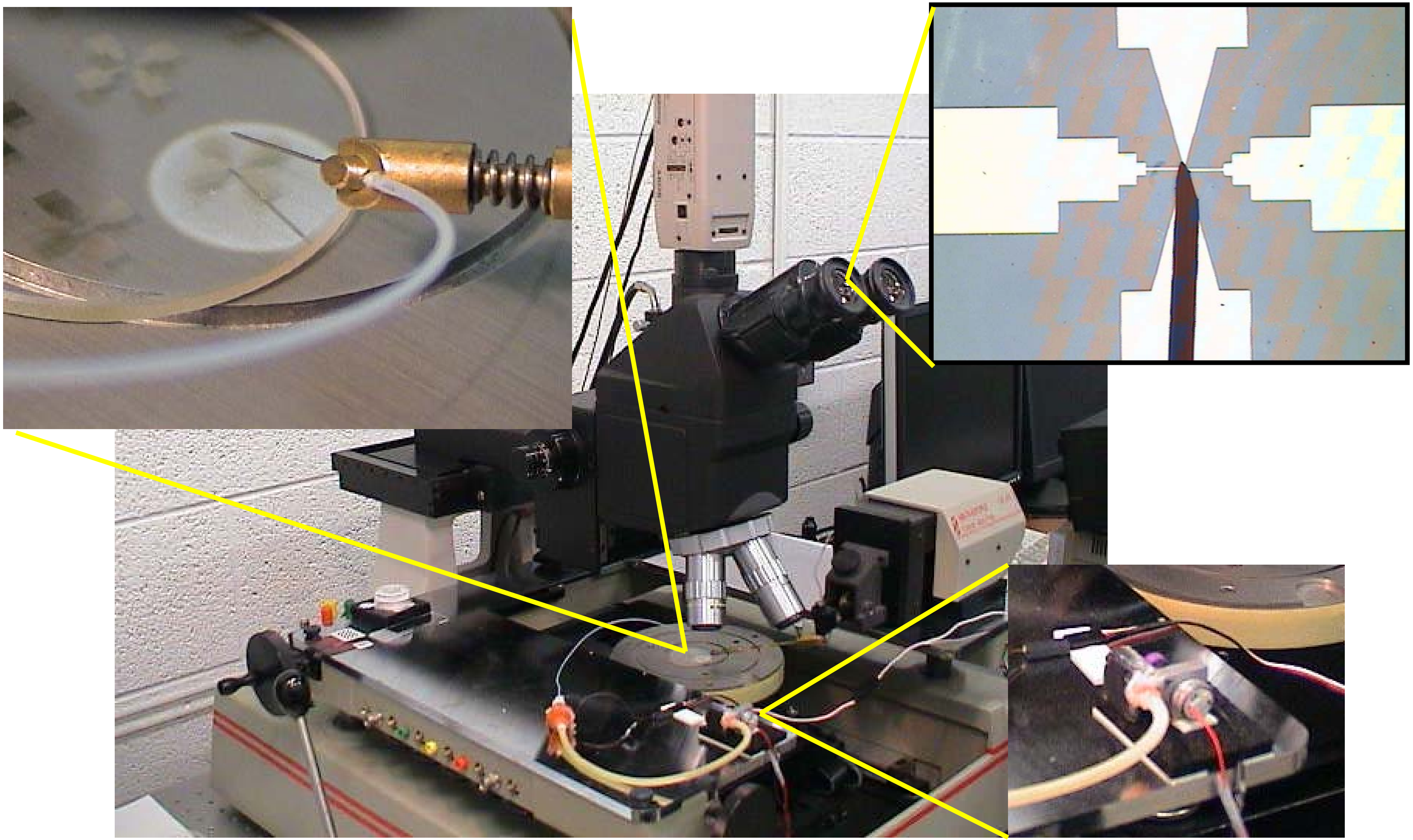


Colonoscopy examines the entire length of the colon; sigmoidoscopy examines only the lower third





CNT Automatic Deposition

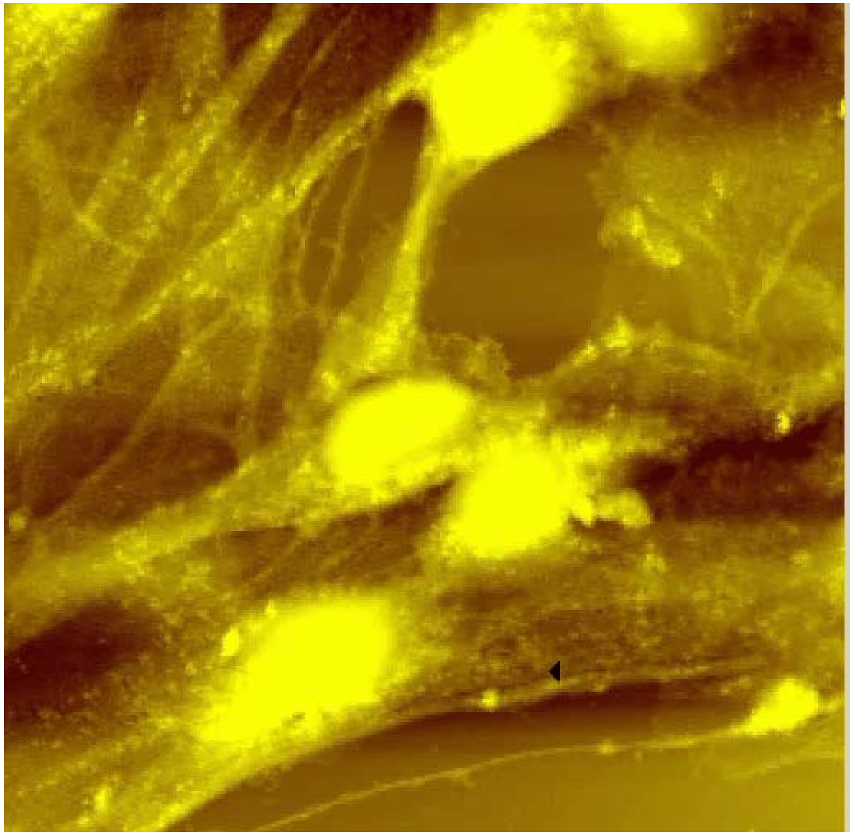


Virus delivery strategy replaced by Trojan horse shuttle paradigm

- Drug delivery system requires targeting: new approach by CNT Trojan shuttle carrier follow by close proximity specific guidance molecules
- Sometime a Nano-Surgically hole boring is advisable at specific receptor site

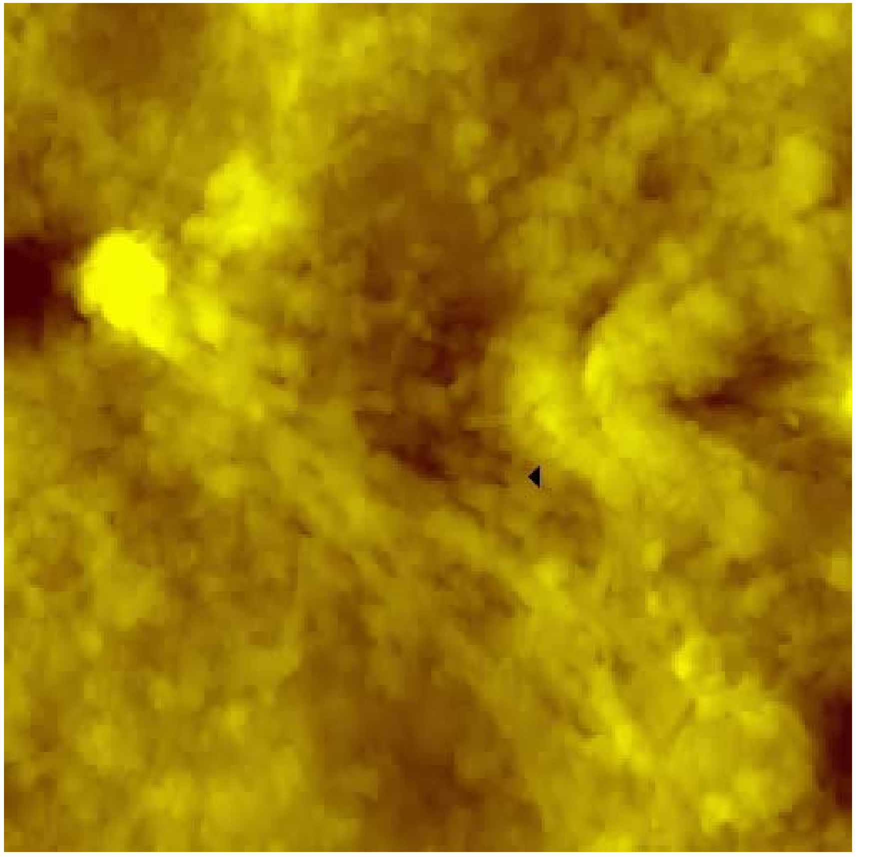
Nano Surgery of Neuron Cell

- The working space is 90μm x 90μm.
- The black triangle is the tip position of the nano robot.
- The objects are living neuron cells cultured on glass surface.
- The goal is to cut the neuron cell axon and dendrite.



Nano Surgery of Neuron Cell

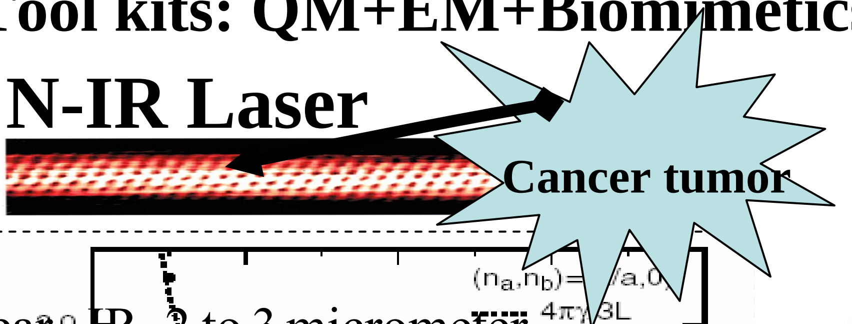
- The working space is 5mm x 5mm.
- The black triangle is the tip position of the nano robot.
- The objects are living neuron cells cultured on glass surface.
- The goal is to dig a hole on the neuron cell membrane surface.

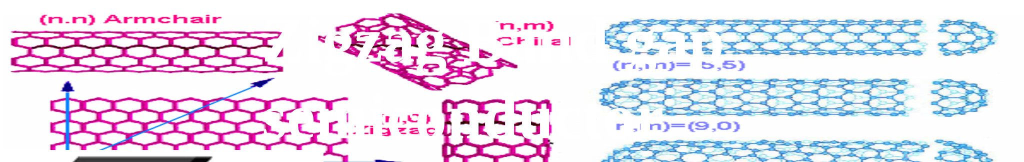


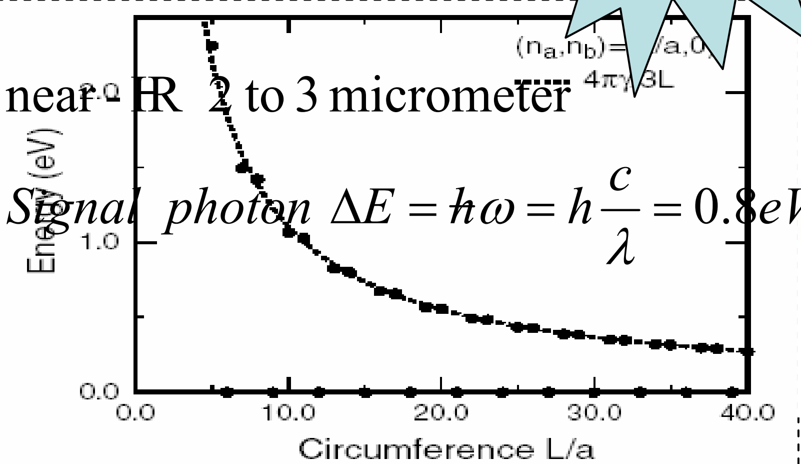
CAD Nanorobot:Drug Proximity Delivery

Nano is not merely sub-micron but QM: tuning of optical energy away from biosystem: Carbon Nano-Tube (CNT), half width of DNA:15 L/a, was guided by Stanford Dr. Dai drug delivery (PNAS Aug 1, 2005) e.g. CNT coated with anionic form of folic acid to go after malign tumor growth vitamin called folate, rapidly absorb near-infrared laser heat to 70^{o}C causing surgical explosion. We broaden CAD Nanorobot at the CNT release site. Broaden CAD Tool kits: QM+EM+Biomimetics called Nanomimetics.

N-IR Laser

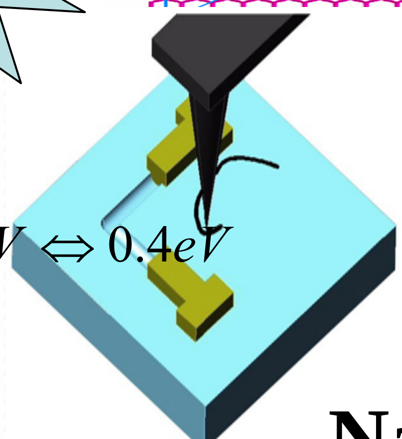






near - IR 2 to 3 micrometer

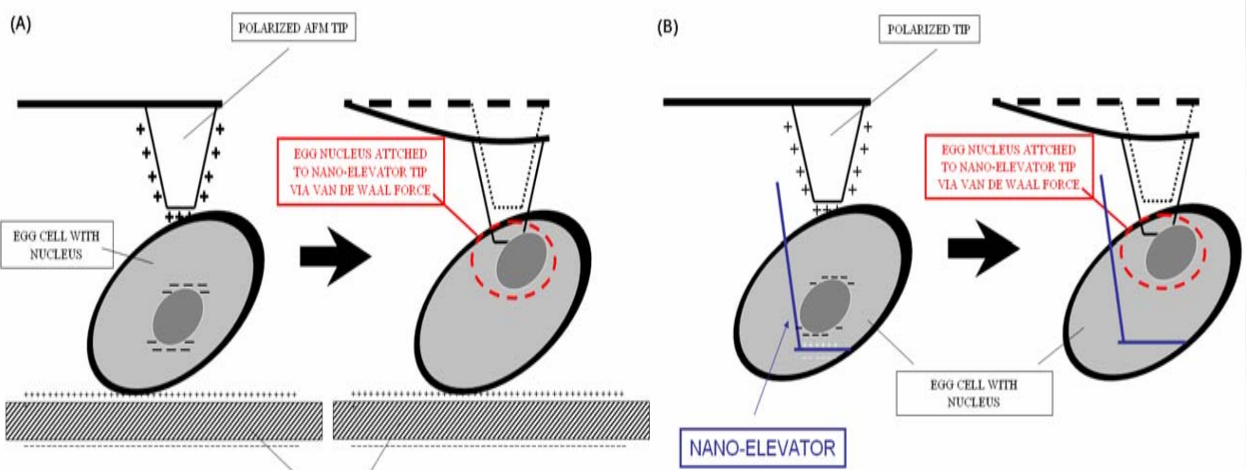
$$Signal\_photon\ \Delta E = \hbar\omega = h\frac{c}{\lambda} = 0.8eV \Leftrightarrow 0.4eV$$

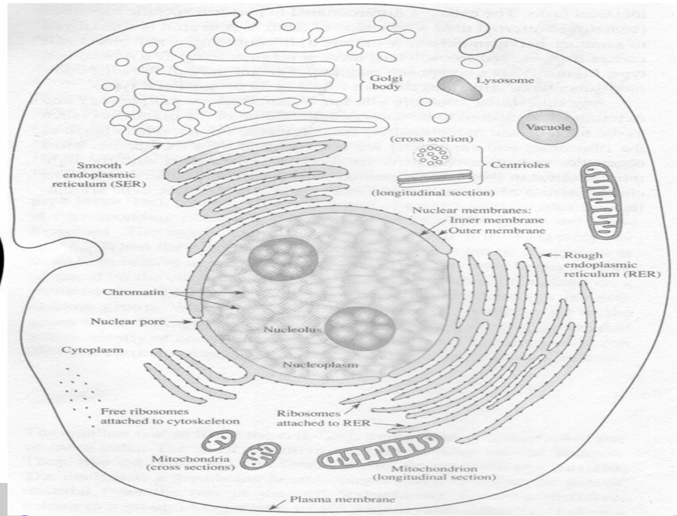


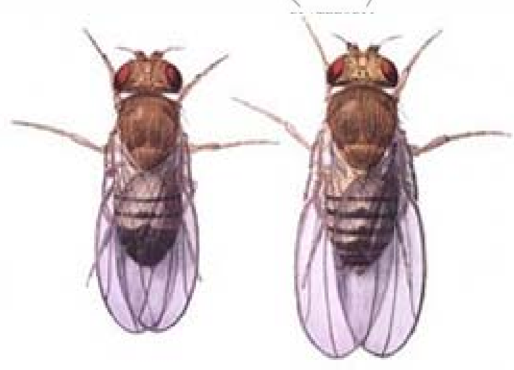
Nano-robot

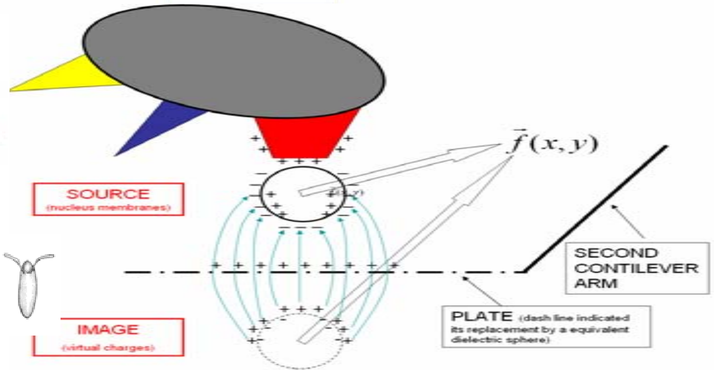
Therapeutic Cloning: (i) Should guest nucleus come servants (**Endoplasmic Reticulum (ER) mitochondria (MT)**)?

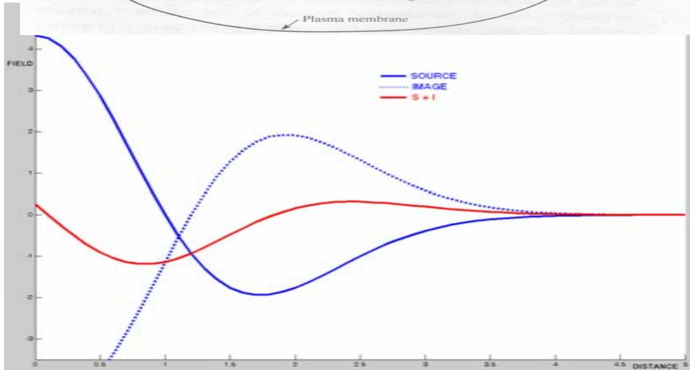
(ii) Can one enlighten the re-programming voodoo procedure?



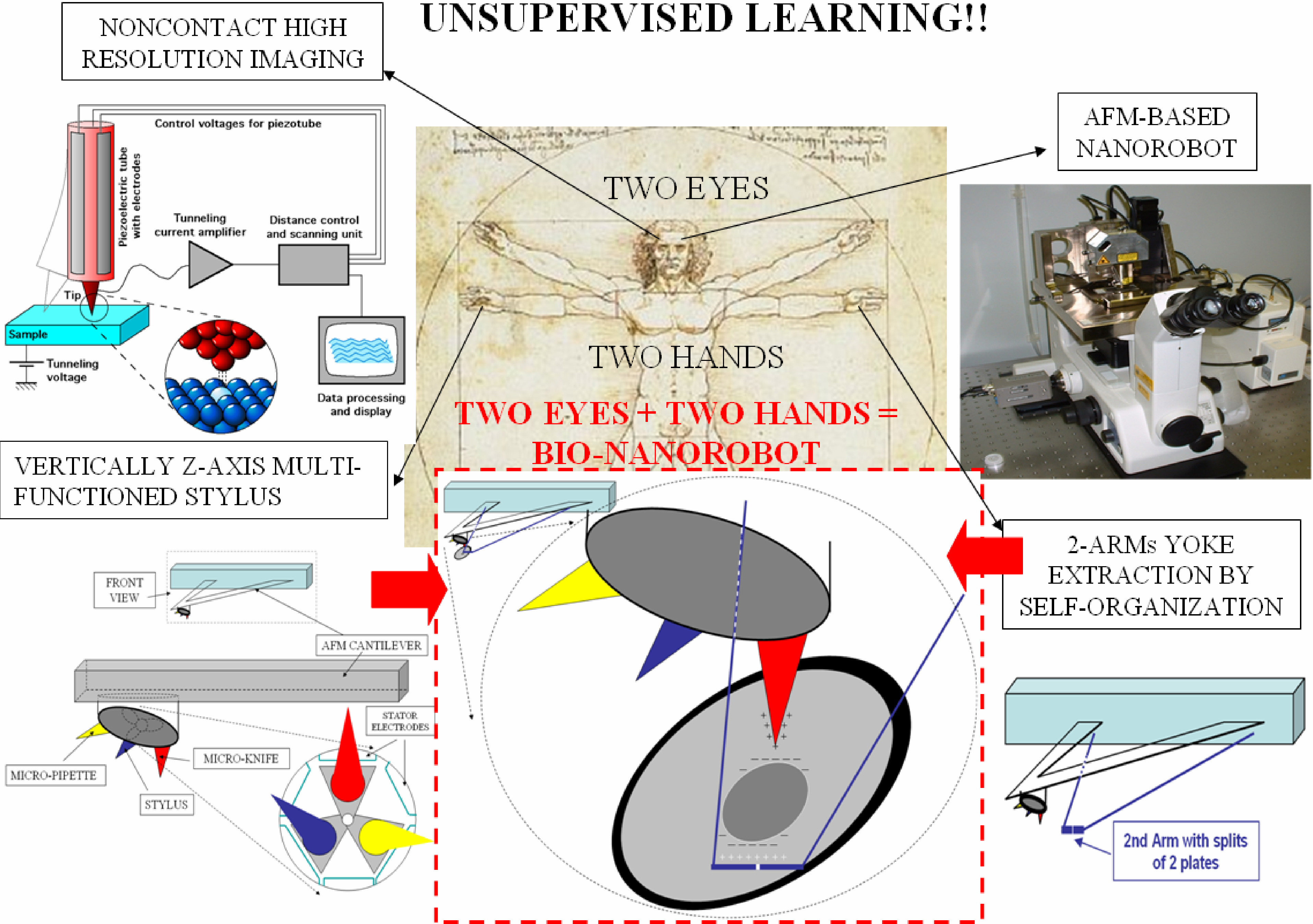




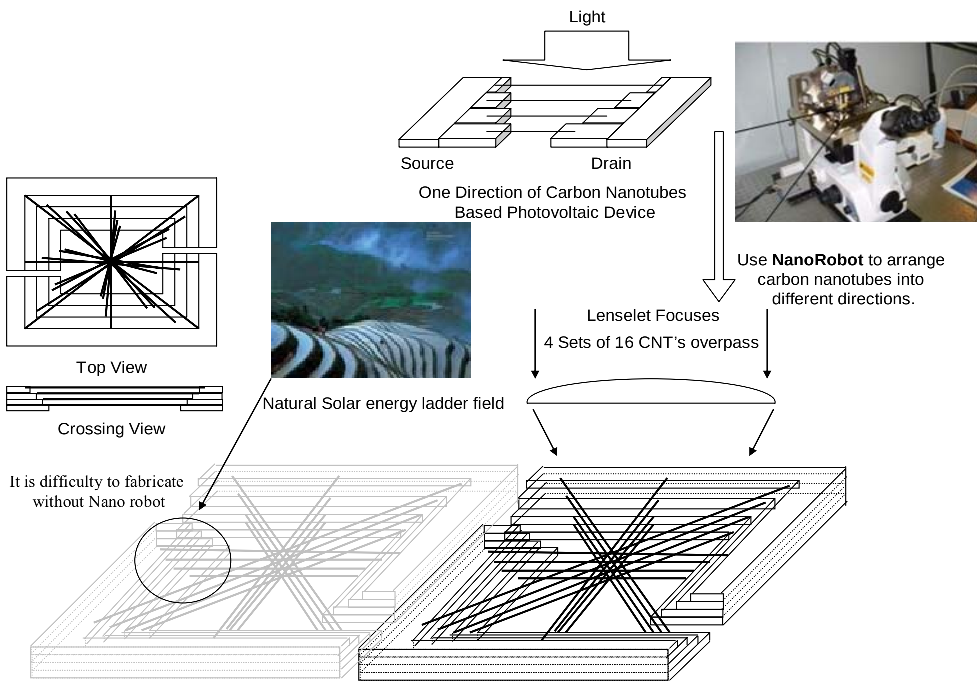


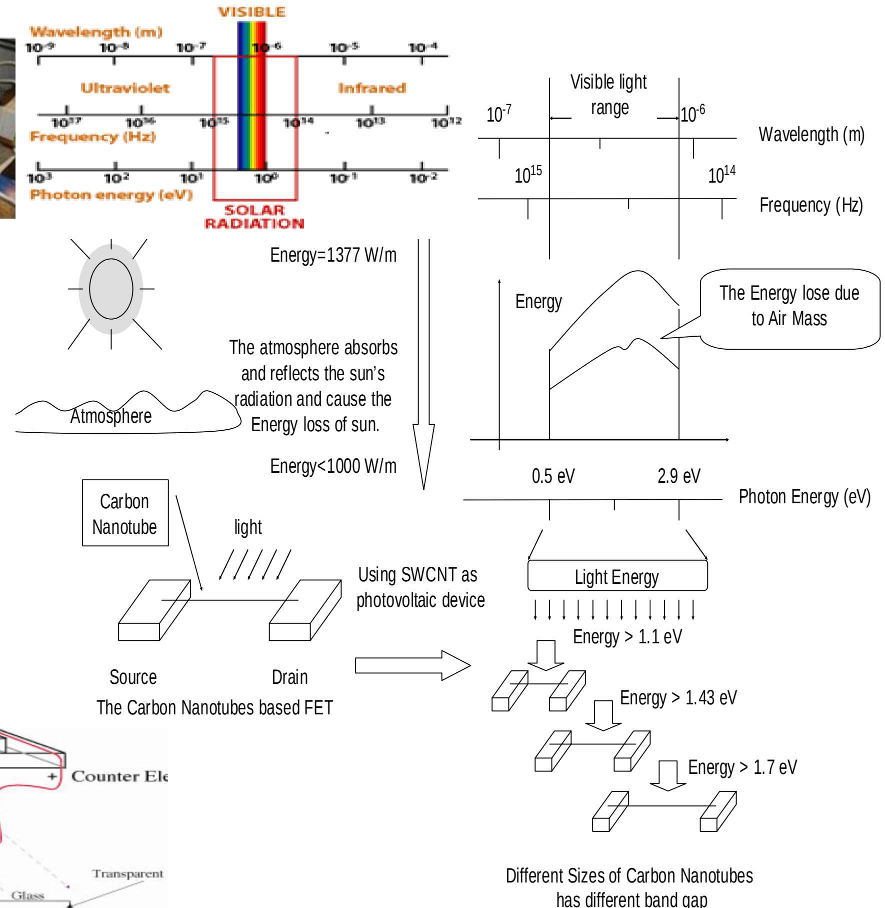


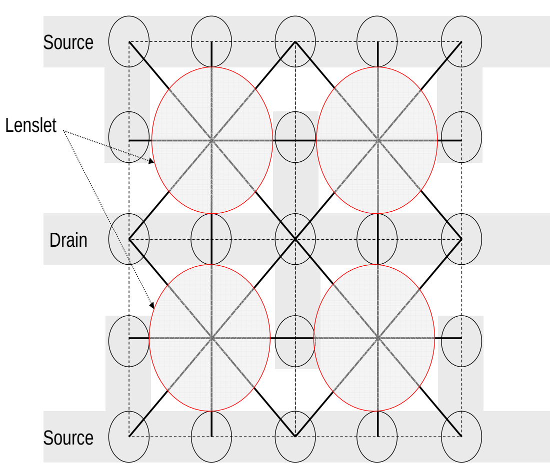
POWER OF THE PAIRs TO ACHIEVE UNSUPERVISED LEARNING!!

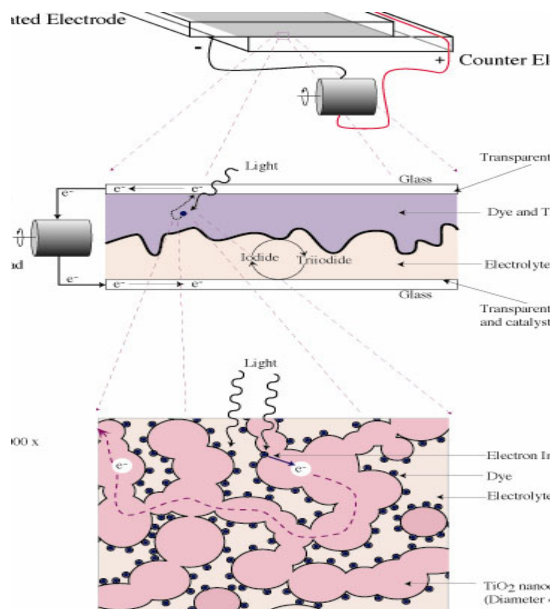


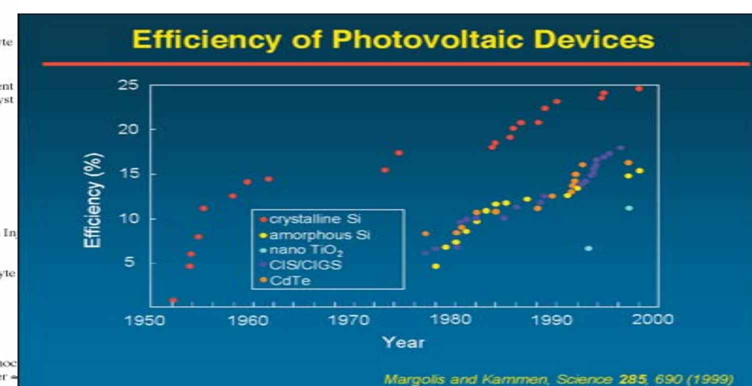
Solar Energy Photo Voltaic Cell











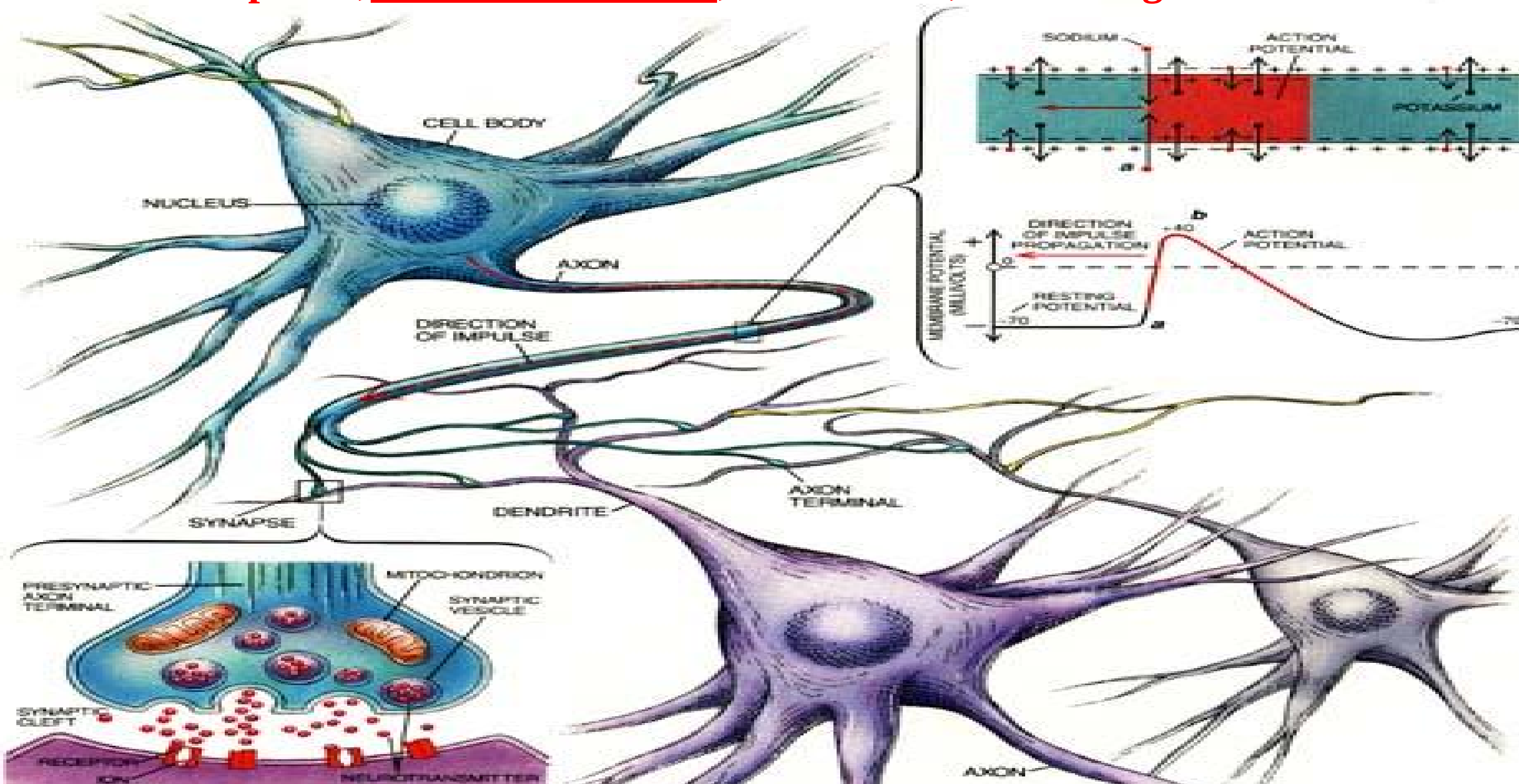
NSF Initiative:*Brain Science* ***Multidiscipline Multi-Millions, Multi-Years***

- March 4-6, 2007, Westin Arlington Gateway Hotel, sponsored by NSF Dir. Arden Bement, Deputy Dir. Kathie Olsen, Chief Scientist, Rae Silver, several different divisions of NSF organized by Chris Wood, Santa Fe Institute divided into 7 groups of 49 scientists.
- Eve Marder (marder@brandeis.edu)) "Nruroethological & Developmental Molecular & Cellular Approaches" need real time in vivo stochastic variability from axon growth path finding from ion channel, receptor, to neurons;
- Partha Mitra (mitra@cshl.edu) "Organization of Behavior Emerging Principle" needs meso-scale net bridging gaps for example the circadian rhythms from molecular genetic, neural, organ, to behavior;
- Jay McCelland (jlm@psych.stanford.edu) "Learning, Plasticity and Development" need timing synchronization coherent engagement, congruity I/O with primitive & connectivity;
- Emery Brown (brown@neurostat.mgh.harvard.edu)" Signal Proc & its development in brain," need neurosci data in different space & time & new chemical modality;
- Jonathan Sweedler(sweedler@scs.uiuc.edu) "Measuring Brian: From synapse to thought," needs biochemical reverse engineering for elucidate the brain functionality imaging;
- Nancy Kanwisher(ngk@mit.edu) "Cognitive Systems across levels of analysis in brain system," requires better modeling and measurement;
- Ted Berger(berger@bmsrs.usc.edu) "biomimetics and the neuron/silicon interface," needs real time bi-direction encode, represent & transform brain science.
- Few would deny major problems facing humanity are social & economic in nature; but few would consider them as the science of 21st Century. Elucidation of human brains might help de-program current radical indoctrination.
- Tools: Internets, Intelligent (Super)computing, Real-time in-vivo Chembio imaging, Nano-technology, Genetic, Epigenetic, System on Chip, Gene chip, Micro-array,

To better understand brain function, advances in available measurement & engineering toolsets are needed

Group of Measuring the brain from synapse to thought

Irv Epstein; Chris Gall; Martha Gillette;Lingjun Li; Anna Lin; Tom Meade; Gordon Shepherd;Jonathan Sweedler; Harold Szu; Mark Wightman



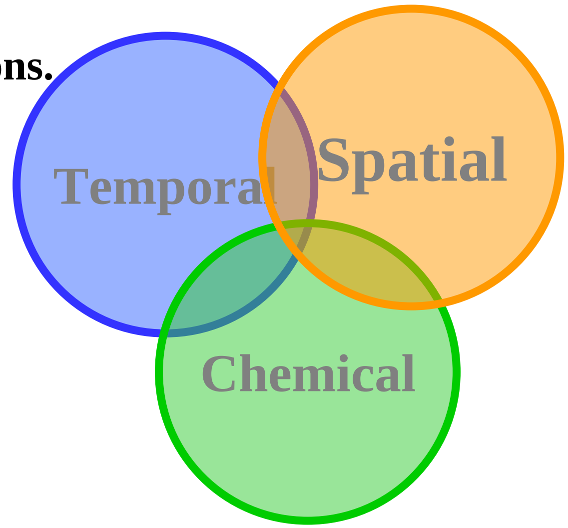
Integration between the ideas of the working groups is crucial

Measuring the Brain: from Synapse to Thought

WHY NOT!:

- Decade of Brain: 1990-2000 (terminated after G. Bush lost re-election)
- Homan Frontier Sciences:1995-2005 (fragmented since US, EU & Asia split),
- 21st Century of Brain 2008-20?? (NSF contemplates with multi-millions, multidiscipline & multi-years deposition of our knowledge in brain science).
- **WHY NOW?**
- **Two Decades of Supercomputing** Centers Nationwide.
- **One Decade of National Nanotech Initiative** e.g. nano-p-imaging, nano-robots.
- **Next Gen Internet**, IP4 to IP16 on line 2008
- **Tools for measuring the brain** can attain integration of **time, length and chemical resolutions in vitro**. Questions remain in hierarchical levels from neurons to neuroethology improving multimodal correlation technologies ***in 3D, in video vivo***.

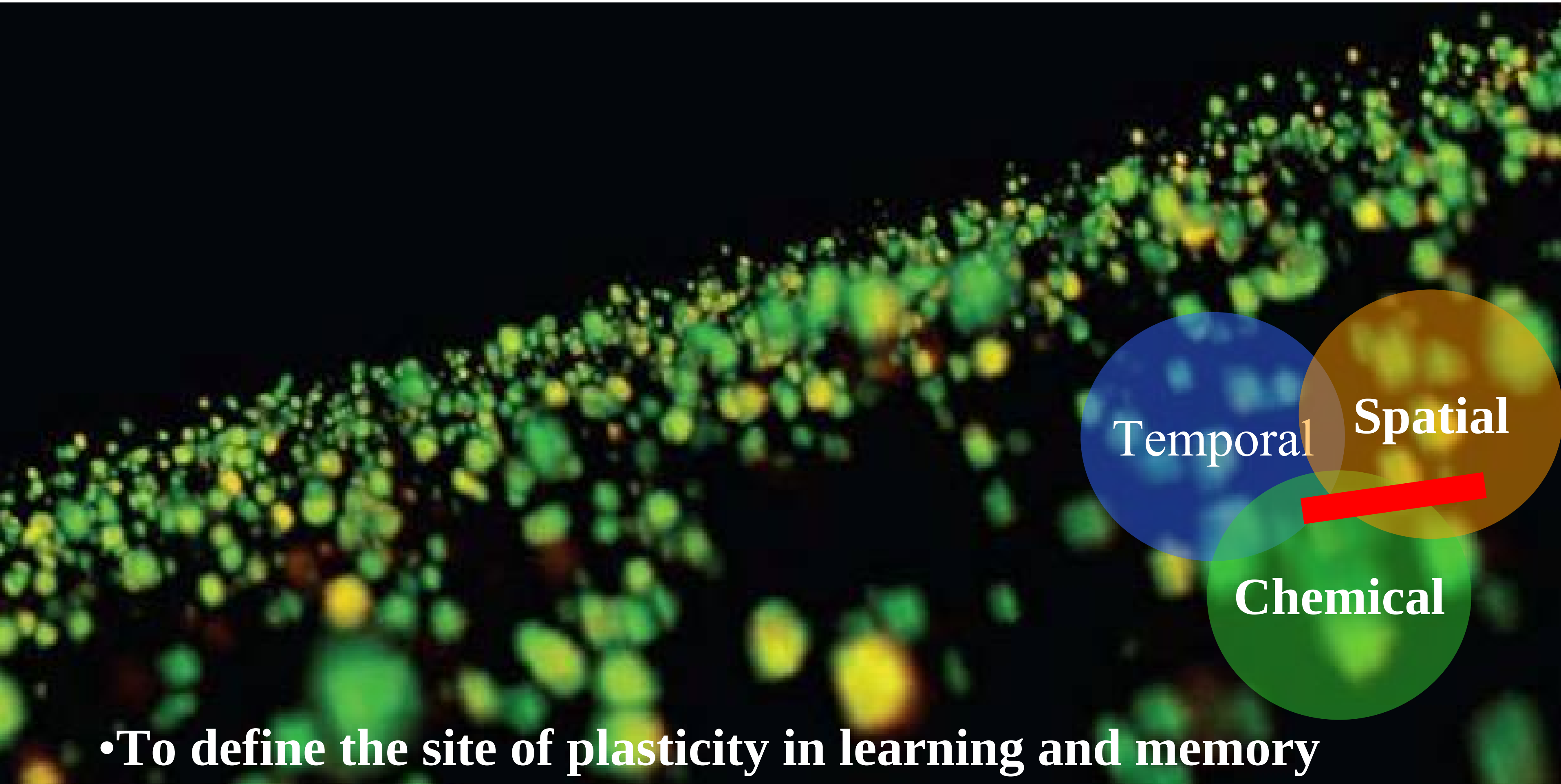
  - **Optical imaging with higher spatial and temporal resolutions.**
  - **Chemical imaging of a greater range of compounds**
  - **Holistic Imaging in whole, living animals**
  - **Multi-level Analysis from organism to subcellular**
  - **Multimodal analysis with coregistration**
  - **Elucidate Intelligence mechanisms by tools**
  - **Systems biology deals with multiple time and length scales**
  - **Atlas of the brain, from RNA, proteins to small molecules**



Measuring the Brain: from Synapse to Thought

WHY NOW? **Tools for measuring the brain have advanced to the point where measurements can be made with time and length resolutions previously unattainable. The future lies with improved multimodal measurement technologies in vivo.**

- **e.g. molecular tagged imaging of plasticity at level of dendritic spines**

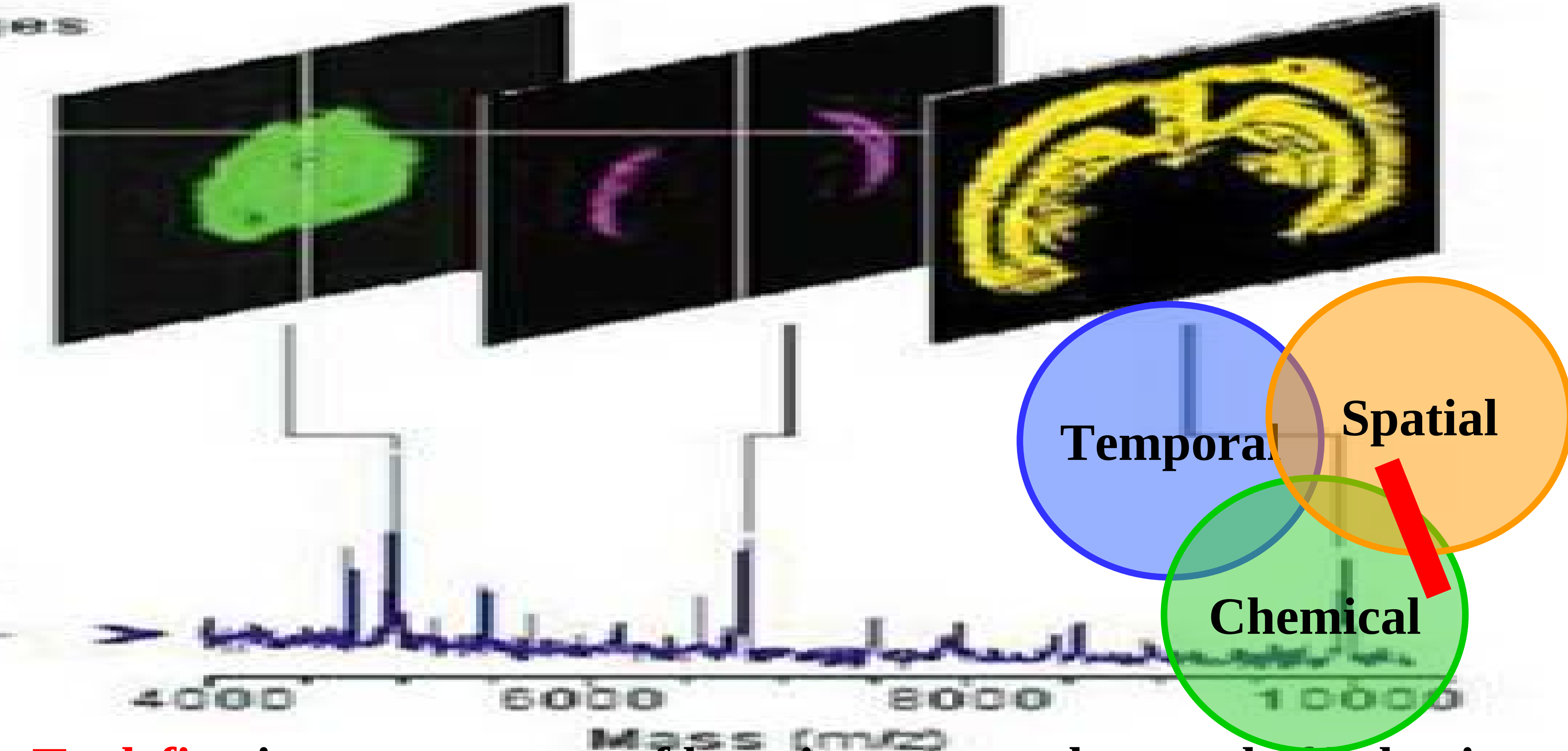


- **To define the site of plasticity in learning and memory**

Measuring the Brain: from Synapse to Thought

WHY NOW? Tools for measuring the brain have advanced to the point where measurements can be made with length and chemical resolutions previously unattainable. The future lies with improved multimodal measurement technologies in vivo in video.

• **e.g. from imaging Mass Spectrometry Chemistry spatial resolution**

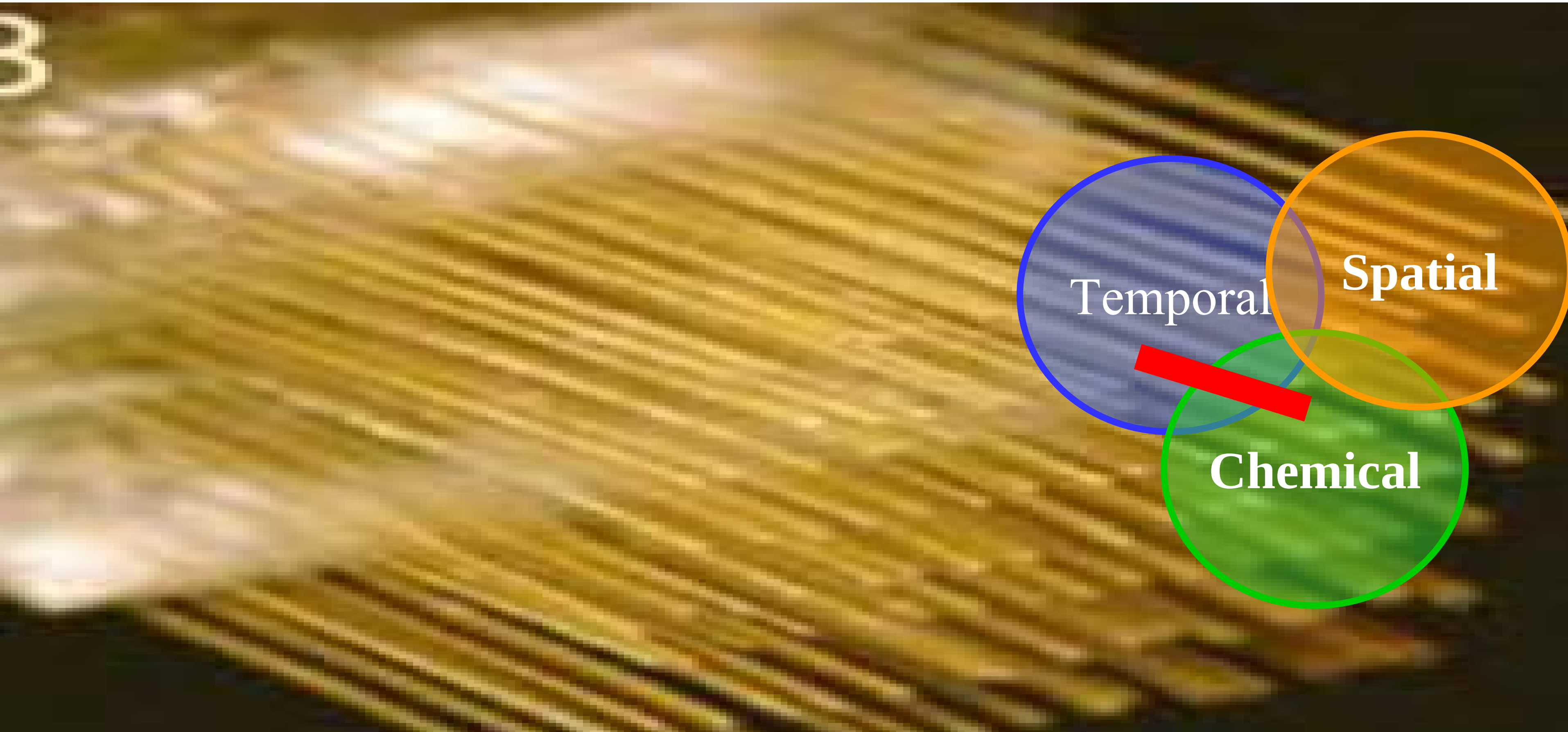


• **To define important parts of learning neuronal control of behavior**

Measuring the Brain: from Synapse to Thought

WHY NOW? Tools for measuring the brain have advanced to the point where measurements can be made with time and chemical resolutions previously unattainable. The future lies with improved multimodal measurement technologies in spatial dimension.

- **e.g. from microelectrodes to sensor arrays for neurotransmitters**

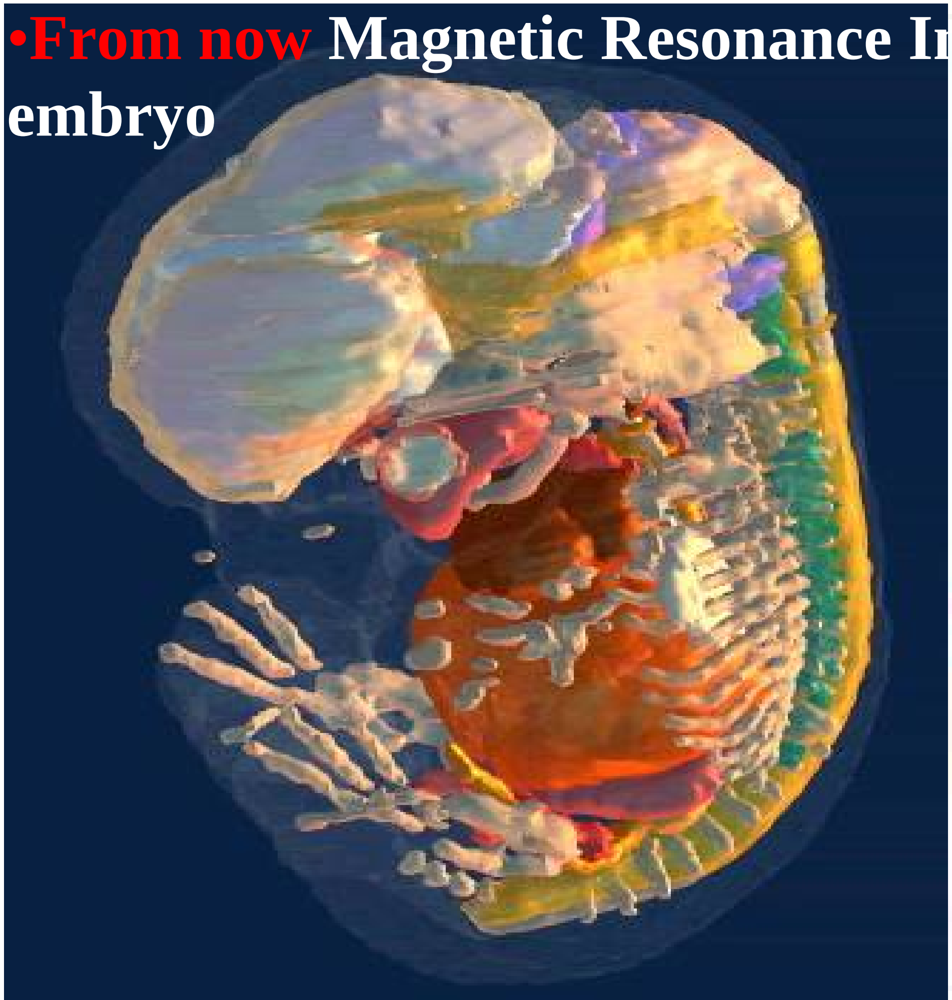


- **To link electrical activity, chemical signaling to understand dynamics**

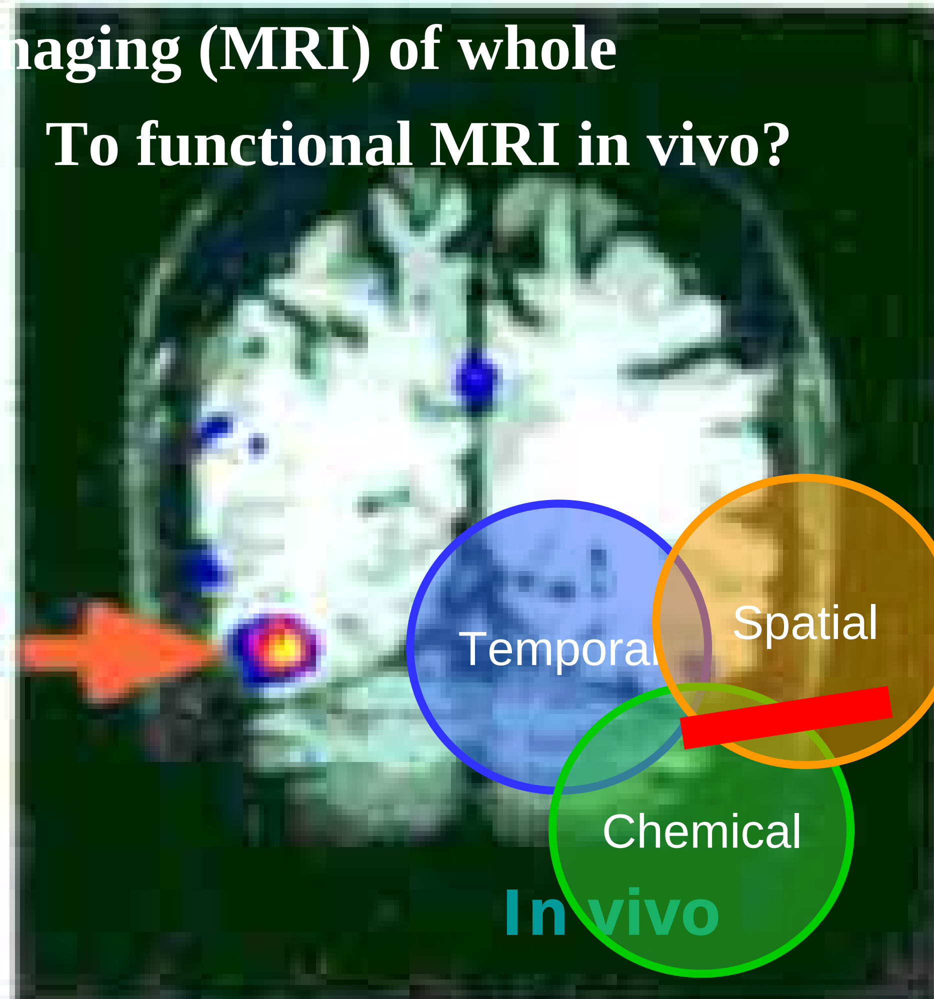
Measuring the Brain: from Synapse to Thought

WHY NOW? Tools for measuring the brain have advanced to the point where measurements can be made with time, length and chemical resolutions previously unattainable. The future lies with improved multimodal measurement technologies in whole organism with higher resolution.

•From now Magnetic Resonance Imaging (MRI) of whole embryo



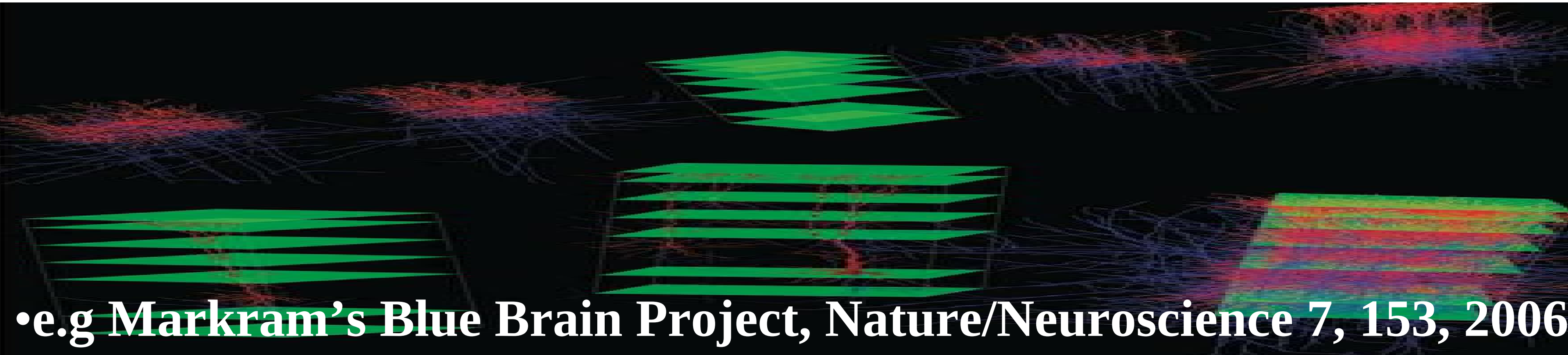
To functional MRI in vivo?



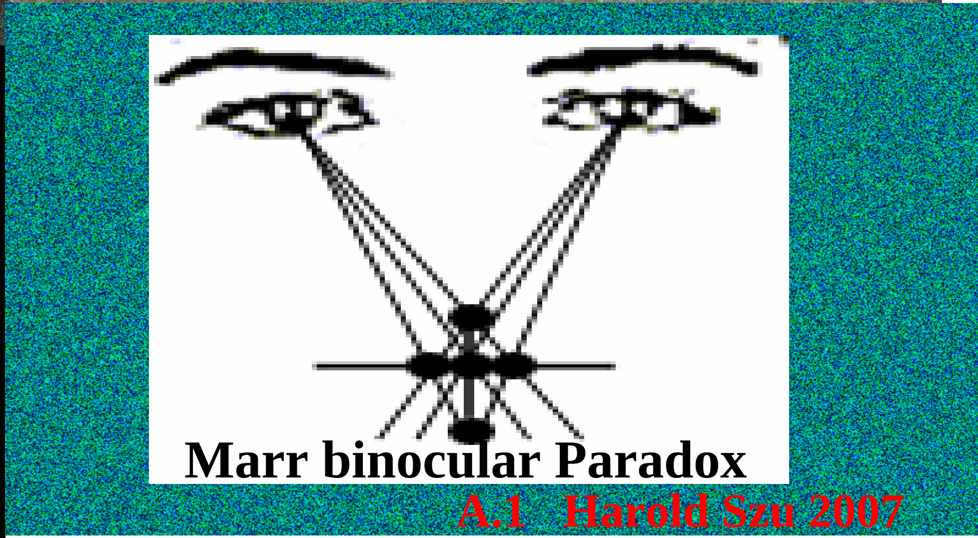
•To elucidate beyond Blood Oxygen Level Difference (BOLD) signal?

How do we analyze, use the data we acquire?

- **Need new math statistical methods to analyze immense data**
  - Filtering out insignificant data?
  - Fusion multiple time and length scales
  - Automatic Pattern Recognition
  - Classification of features into classes and comparison
  - Interpolation & Extrapolation of incomplete data sets
- **Need data sharing and archiving at NGI/Supercomputing Centers**
  - At the national level to support experiments and simulations
  - Long term commitment needed to replace Human Brain Project
  - Hierarchical levels of data needed
  - Integrate across different animals:invertebrates & vertebrates



Impact Area: Why do we have pairs, e.g. two eyes, etc. ten sensors for five sense inputs? **A reason is "*agree, the info;disagree, the clutter*" at the constant 37^{o}C brain temperature for a soft-decision of agreement for instantaneous response for the survival of fittest species (Unsupervised BSS)**



Educational opportunities

It's the People, People, and People!

Promote cross-training across disciplines in multiple tenures

While NSF training opportunities exist, they tend to be divisional or disciplinary and locked-in.

How to remove barriers & level the play field

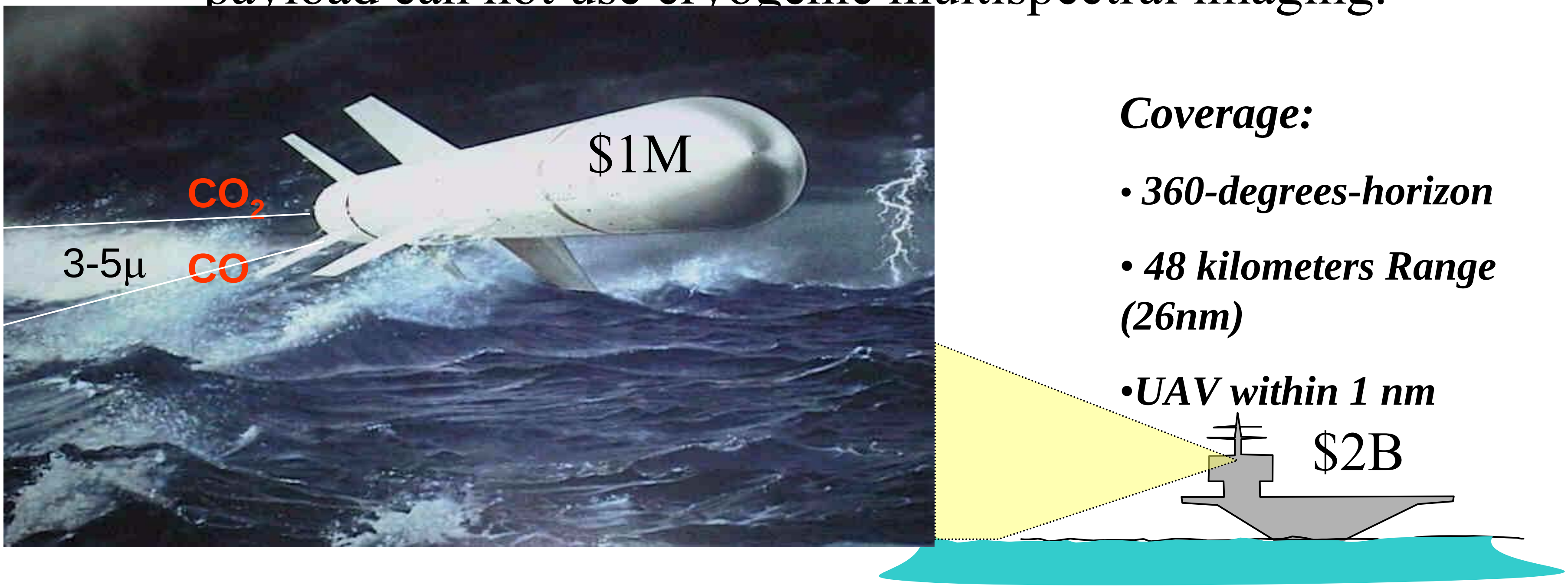
•Graduate student training plans outside of centers or defined plans

•Postdoctoral training across fields

•Release time for faculty to study in a new discipline

How to increase Stand-off Distance of Missile Defense?

In-situ mini-UAV can give OTH carrier's scouting needs, but the payload can not use cryogenic multispectral imaging.



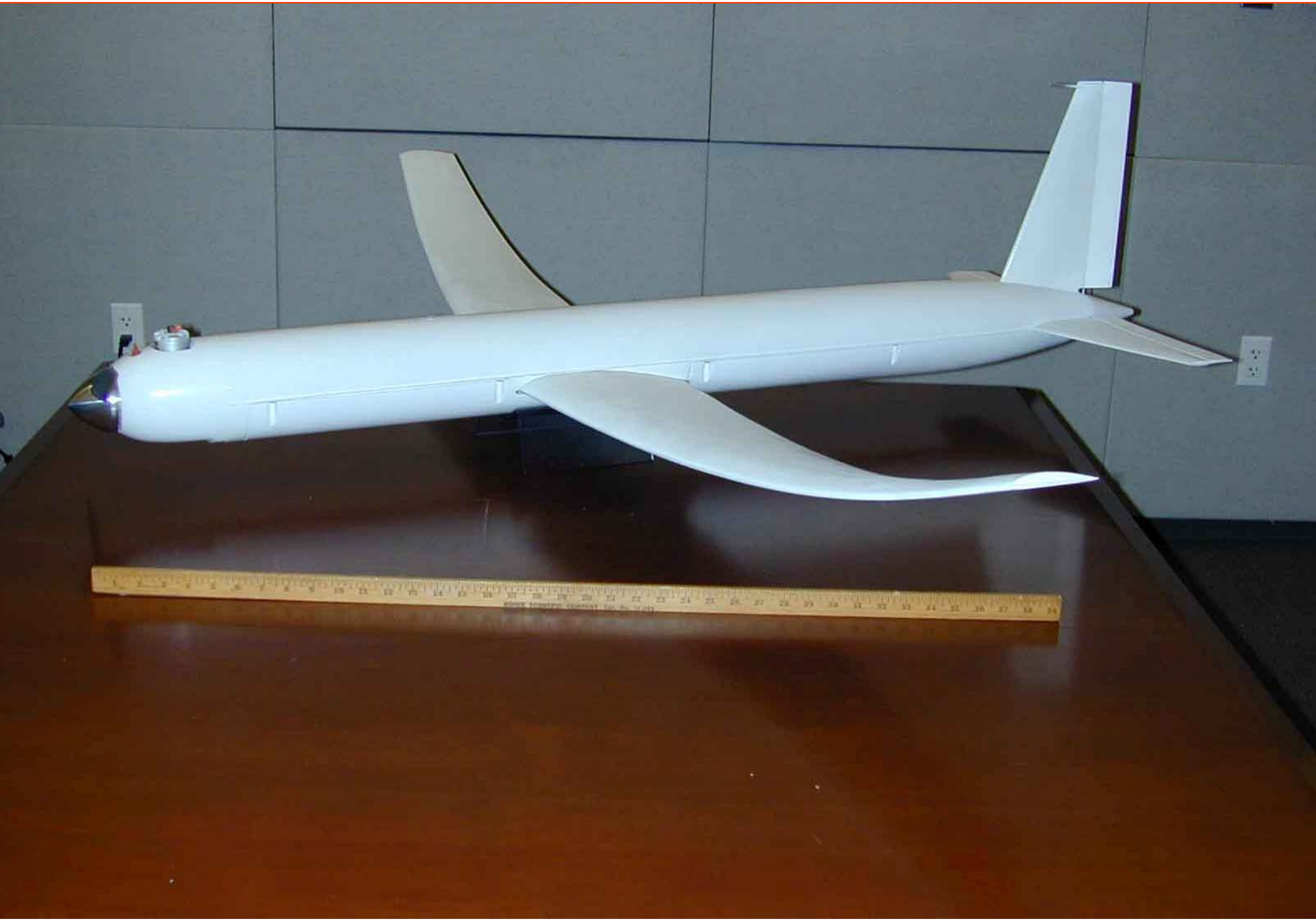
Coverage:

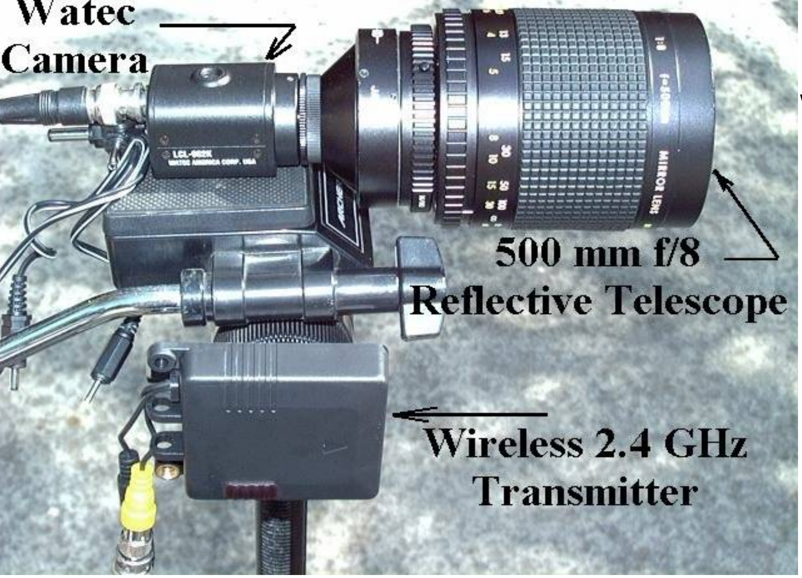
- ***360-degrees-horizon***
- ***48 kilometers Range (26nm)***
- ***UAV within 1 nm***

•UAV can steer itself by means of Long IR FPA to point at incoming cruise missile plume looking for unique feature at Mach cone turbulent mixing range dependence of CO-CO_2 Mid-IR spectral lines.

•Thus, FNC platform protection can increase the standoff distance by unsupervised fusion in terms of two color IR by UAV as X-band Radar can not yet see OTH missile due to the earth curvature ocean waves.

Platform Resource Constraints: no gimbals mounting for EOIR-RF



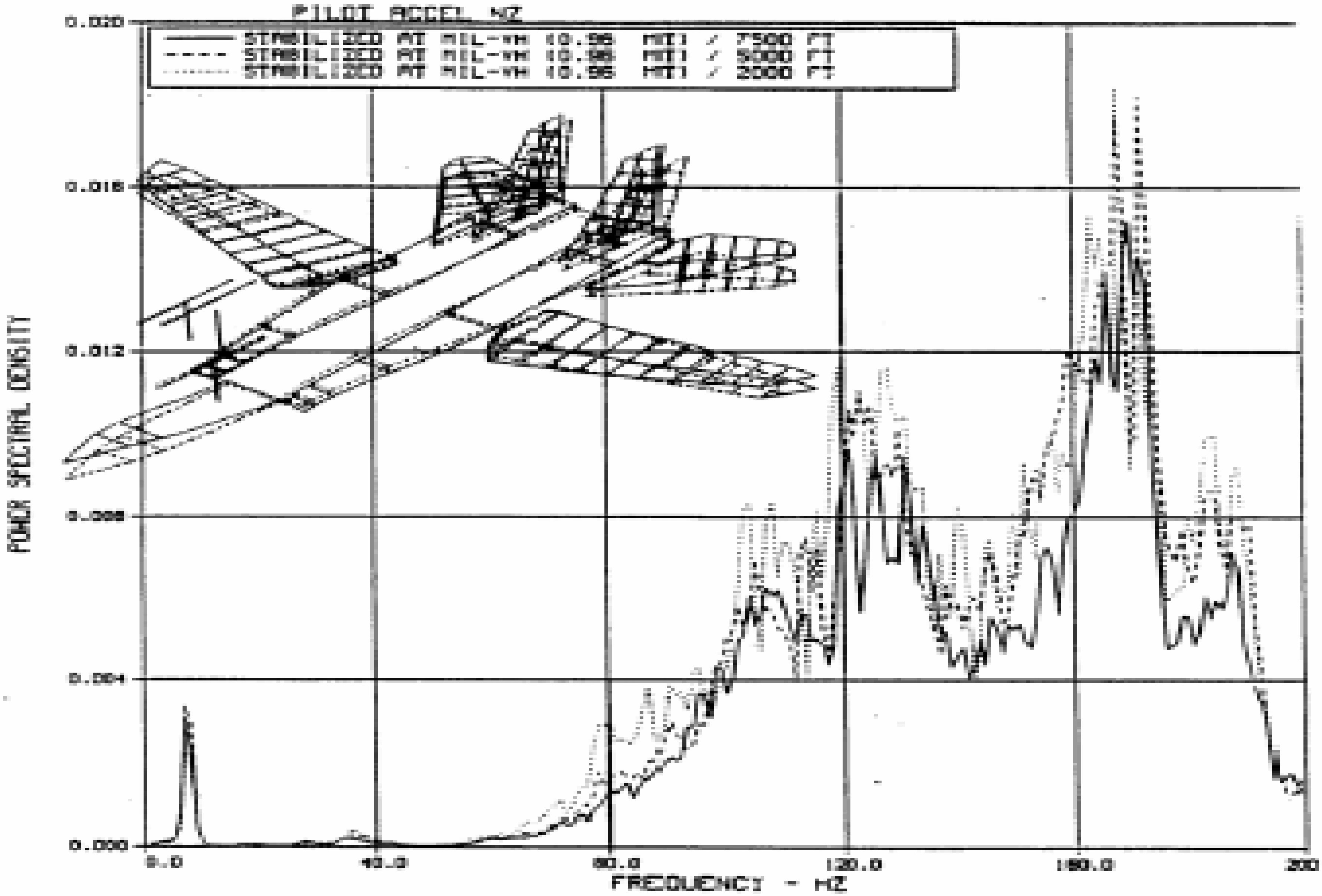


Why video sub-pixel jitter correction is important?

•We define jitter to be sub-pixel or small-amplitude vibrations up to one pixel, as opposed to motion blur over several pixels for which there already exists real time correction algorithms used on other platforms.

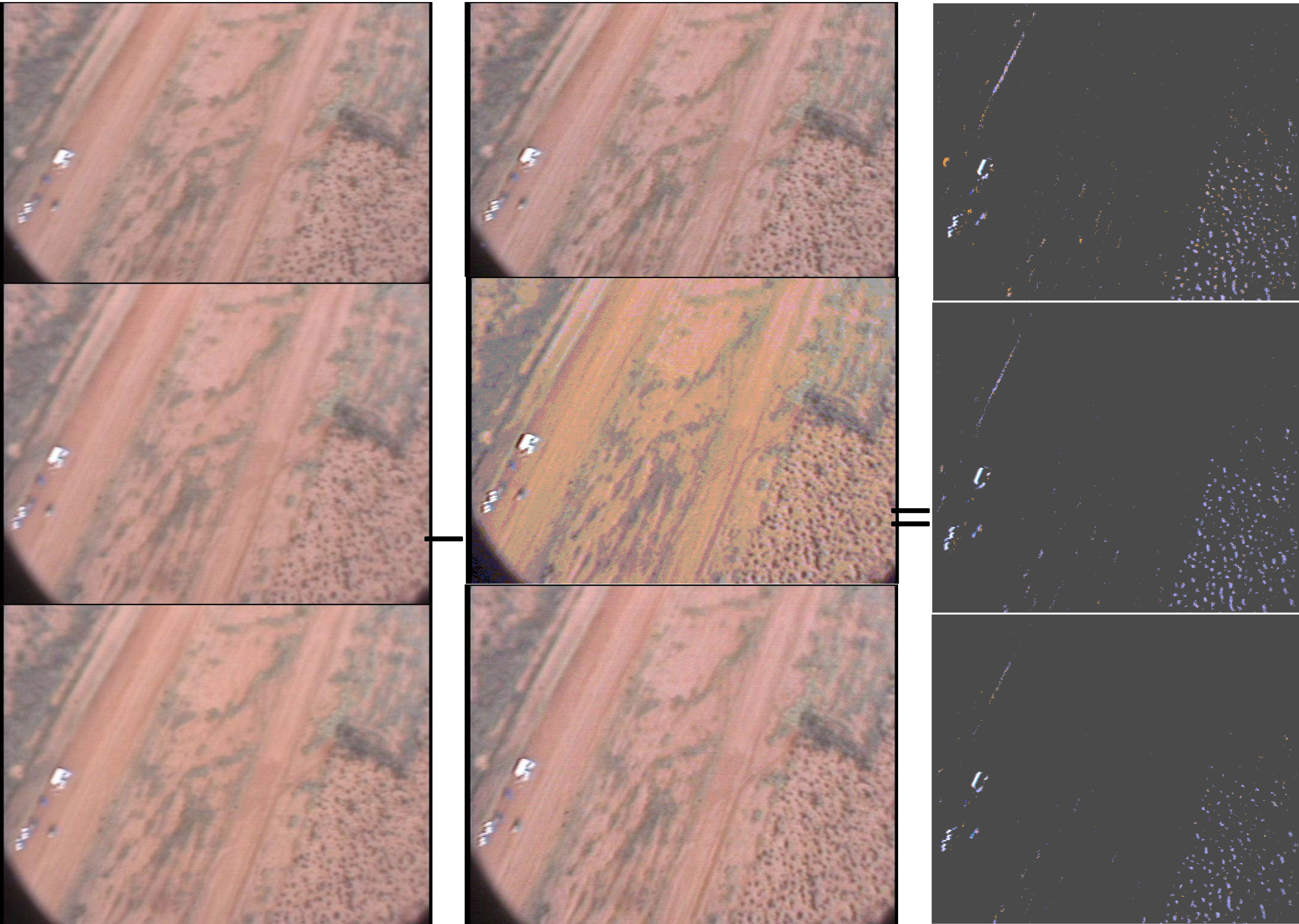
•Since micro-UAV, Silver Fox, cannot afford Gimbals mounting from the isolation coupling to the turbulent aerodynamics of the airframe, we must explore **real-time unsupervised learning software** on board of μ-UAV to mitigate the sub-pixel jitter effect.

•The sub-pixel accuracy is the basis of affine distortion transform and passive cell-phone transceiver array one per UAV for interference SAR registration.



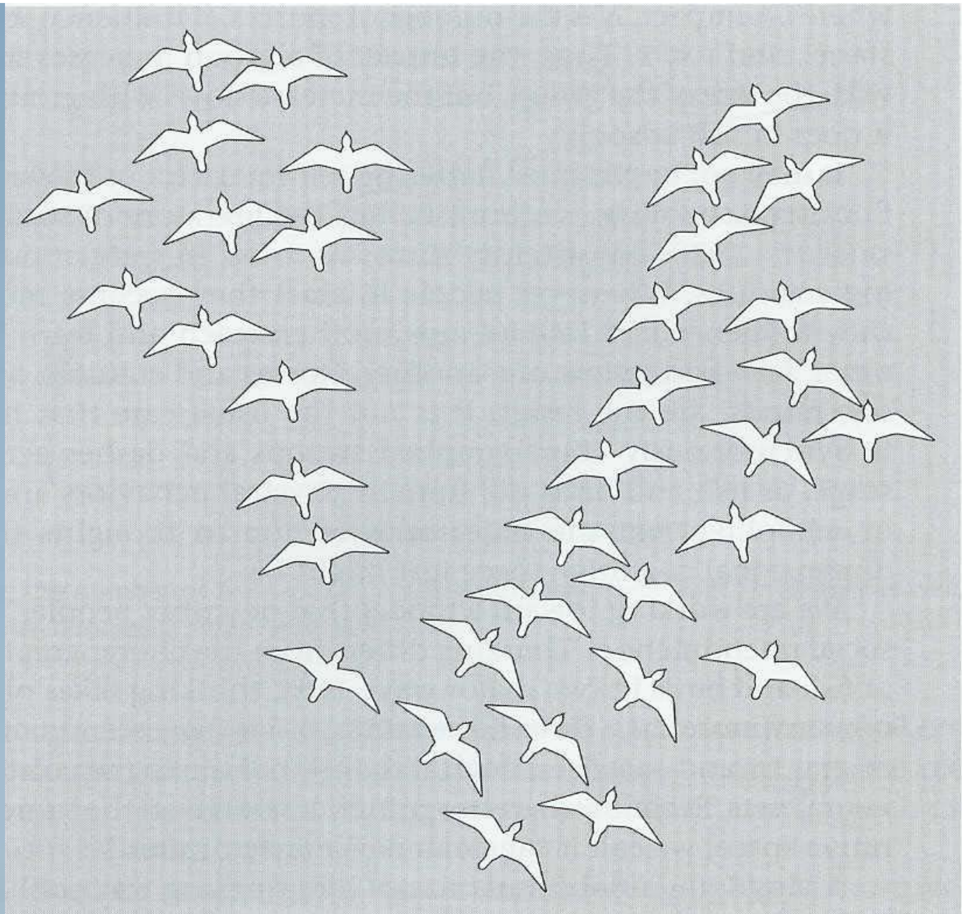
From pg 30 & 40 of
Grumman Aerospace Corp., Report No. LD-303D-89-002

m-UAV sub-pixel jitter Algorithm by Szu: ***"agree, signal; not, noise":***

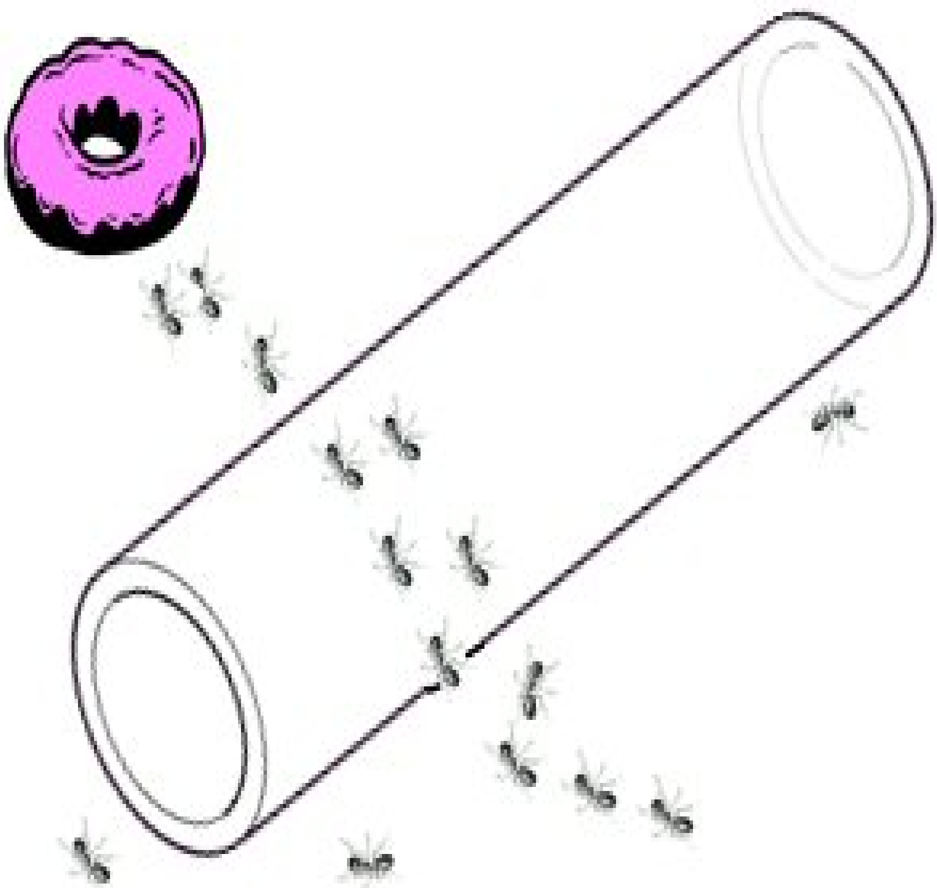


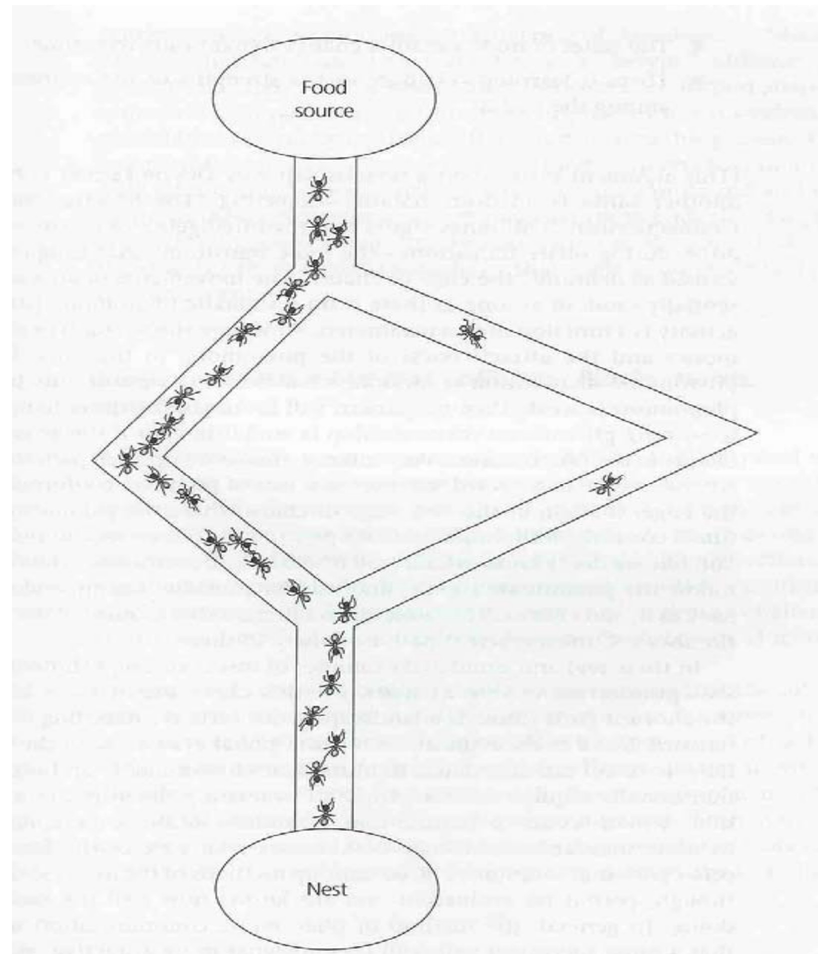
Swarming Intelligence: What is the minimum communication of each node required to make collective behavior more intelligent?

The swarm intelligence in bird migration is based on three interaction principles: moving to center, maintaining the same speed, and avoiding collisions.

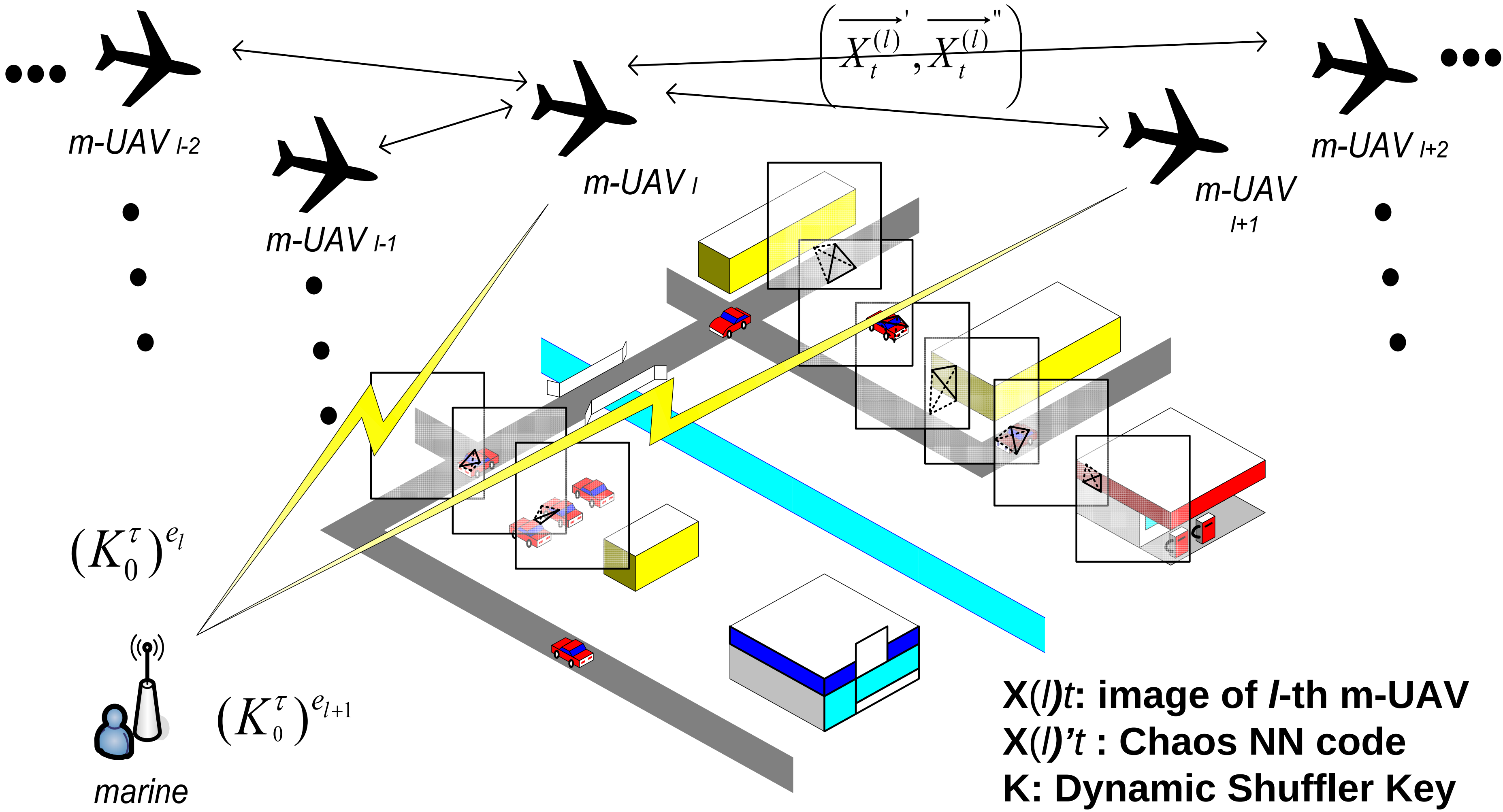


Ant's solve TSP because pheromones accumulates more on shorter path since ant sets out on that returns faster





Robust Private Chainsaw Communication
A minimum of 5 m-UAVs requires security; but "no man, no NSA"
Jittering Mosaic Image Processing (Szu et al. SPIE ICA etc 2006)



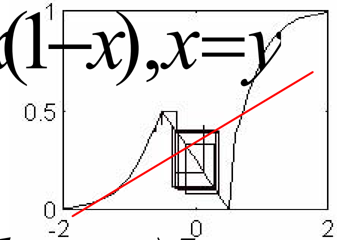
X(l)t: image of l-th m-UAV
X(l)'t : Chaos NN code
K: Dynamic Shuffler Key
X(l)t": hashed chip

Pair-wise Privacy by Chaotic Neural Nets

Binary Message

Carrier of Message

$$y=f(x)=4\lambda x(1-x),\ x=y$$



$$M' = EXOR\,[\,M\,, Chaos\,] = Mod_2\,[\,B(M) \oplus B(Chaos)\,]$$

$$M = EXOR\,[\,M'\,, Chaos\,] = Mod_2\,[\,B(M') \oplus B(Chaos)\,]$$

Rivets, Shamir, Adelman (RSA) codec N = pxq two primes; e.g., 15 = 3 x 5; (1)

(I) x (p-1)x(q-1) + 1 = e x d; public x private; 1 x (3-1) x (5-1) +1 = 3 x 3; (2)

Coded message M'= L = remainder [M^d/N]; (3)

Decoded retrieval M = remainder [L^e/N]. (4)

For example: M= "Chaos typing & initial seed is on page 3"

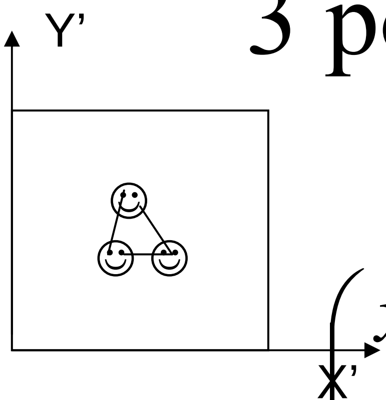
Given a message, say ASCII "C" = 3;

Coded M'=L= remainder [3^3/15] = [27/15] =12;

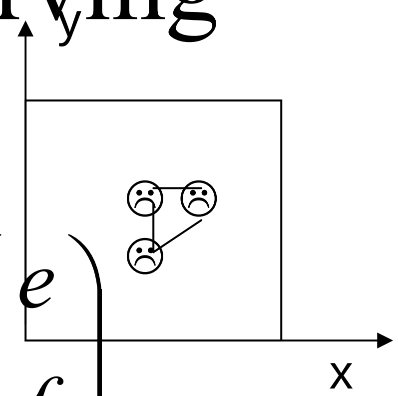
Decoded M = remainder [12^3/15] = [1728/15] = [115+3/15] =3.

Chinese Remainder Theorem & Square & Multiplier Lookup Table are used. QED.

Cooperating Surveillance distortion registration by simple Affine Transform 3 points area-preserving



$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix}$$



Scatters of inexpensive plastic disks having both optical & thermal signatures provided the required minimum three correspondent points in neighborhood frames (over-determined case is given by Szu 1980). Then, location of approximate centers of scatters is geo-registered.