DETERMINED BLIND SOURCE SEPARATION USING NOVEL MACHINE INTELLIGENCE

Names

Denis Ombati

Supervisors

Dr E. N. Ndung'u and Dr L. M. Ngoo

Masters of Science Thesis research progress

Telecommunication and Information Engineering (JKUAT)

2014

OUTLINE 1:1 OF

- Background and Motivation
- Problem
- Literature
- Design of the BSS Model and Methodology
- Scope of the work done to Completion.

Problem Statement 1: 2 of 6

- A hands-free speech recognition system and a hands-free telecommunication system are essential for realizing an unconstrained, and stress-free human-machine interface. In real acoustic environments, however, the speech recognition performance and speech recording performance are significantly degraded because one cannot detect the user's speech with a reasonable signal strength (high *signal-to-noise ratio (SNR))* owing to the interference signals such as noise.
- A blind speech processing, free from prior weakness, is particularly in demand in our communication systems.
- Such process should be able to adapt to different conditions; Linear and Non-linear.

Goals and Objectives I: 3 of

Main Objective

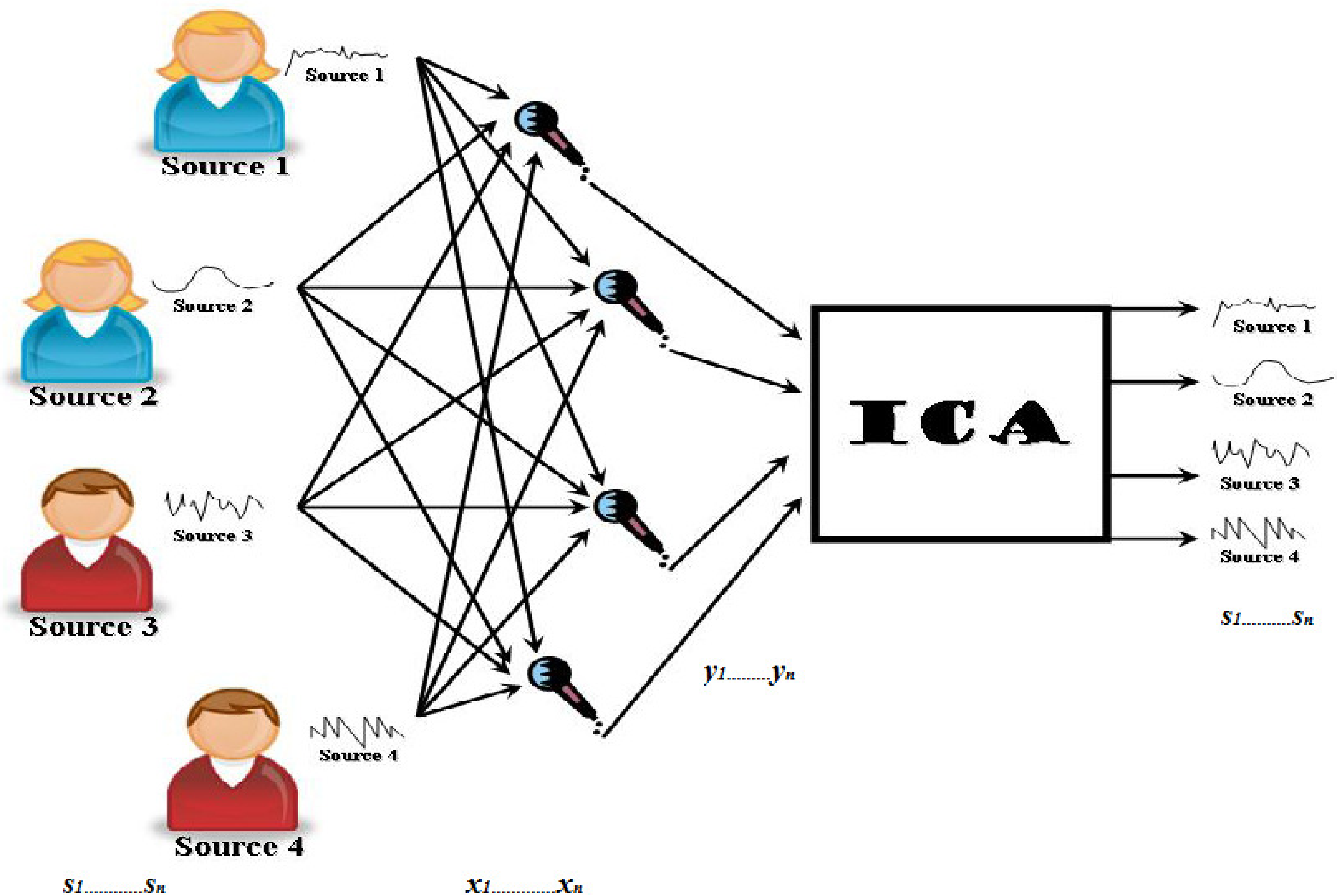
- The main objective of this paper is to simulate BSS system model using Independent Component Analysis and optimized with Radial Basis Function network ICA-RBF. The system will be based on the output sources signals, sensor signals as well as some prior knowledge of the mixing system.

Specific Objectives

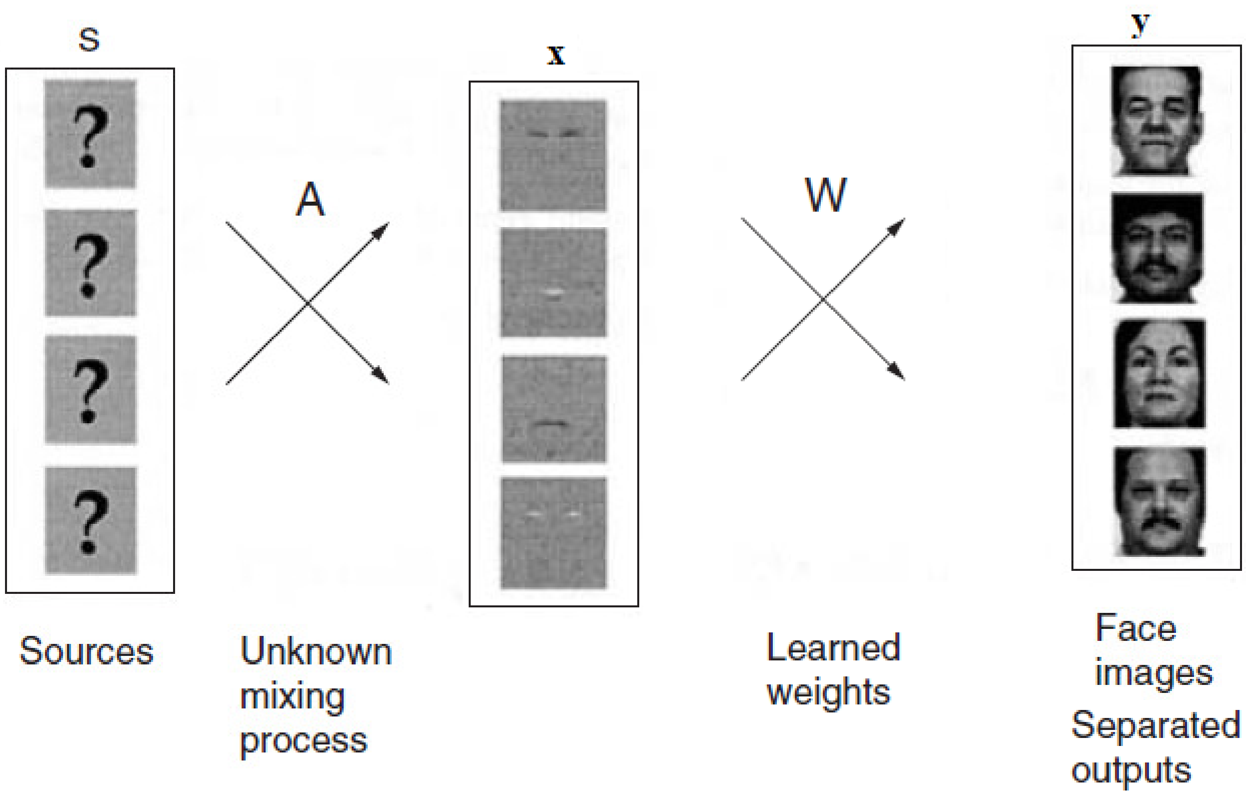
- Analysis of ICA methods based on Information theory using statistical SP.
- To simulate the performance of neural based BSS artificial neural network (ANN) model for optimal results by harnessing the power and parallel structure of neural networks and multi-core processors. A property proposed to suit multi-channel blind source separation system.
- Compare the performance of the proposed technique based on signal separation strength ratios with reference to already popular BSS method and then draw conclusions on proposed ICA-RBF Network system based on the attained performance indices.

INTELLIGENT SYSTEMS IN "COCK TAIL PARTY PROBLEM"

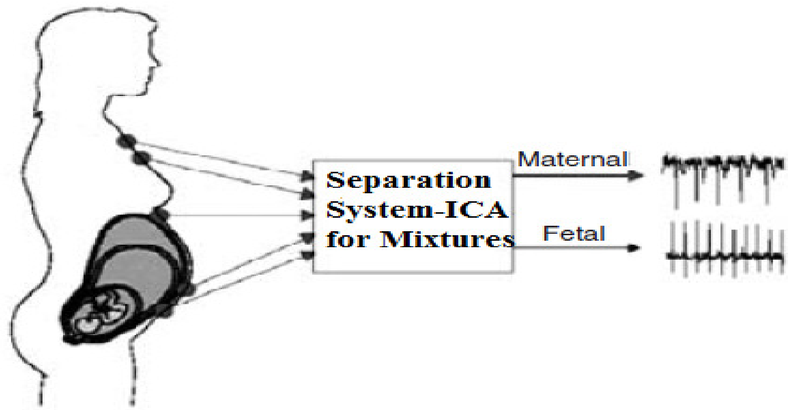
I: 4 OF

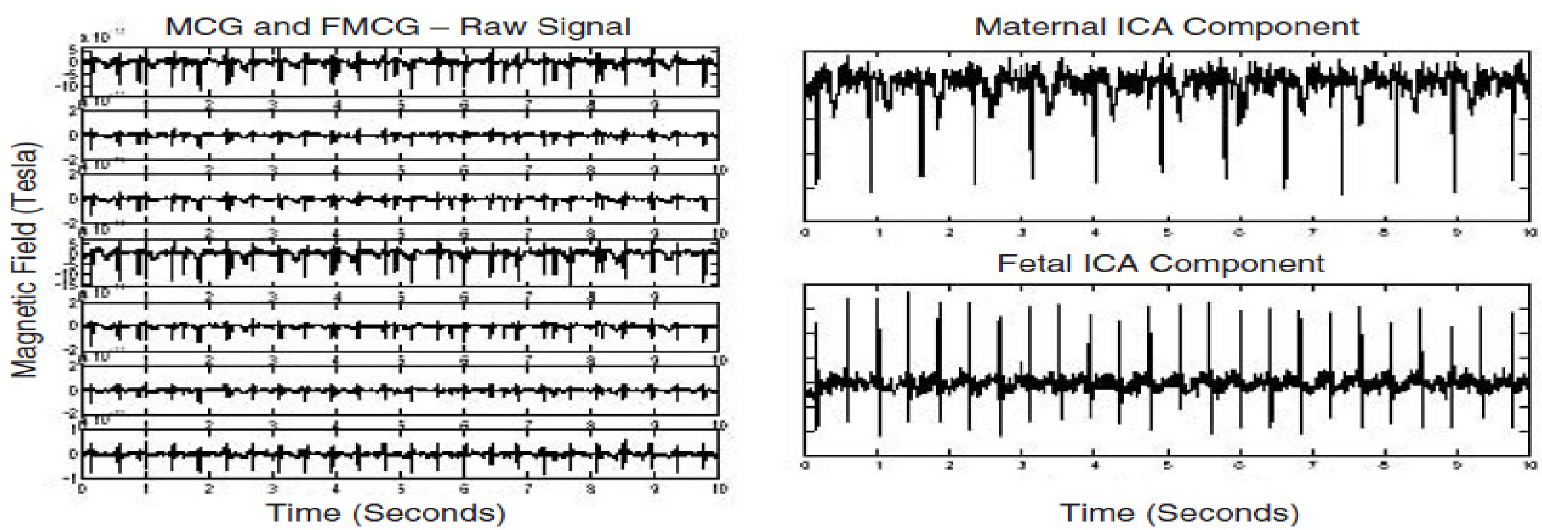


Introduction to Blind systems. 1:5 of



Independent Component Analysis ICA in medical field 1:6 of





Justification of the study

- Blind signal processing is one of the emerging researchable areas in digital signal processing (DSP). Diverse application of BSS include recording and relaying multiple speech sources, satellite-mobile systems, and medical diagnosis of fetal heart rate (FHR), and maternal electrocardiogram (MECG) in fetal electrocardiogram extraction (FECG).
- Especially in Teleconferencing there have been efforts put in place to improve the acoustic environment to satisfy the users, but generally, these environments are acoustically challenging with lots of hard surfaces like plaster, glass, wooden walls, hard ceilings, boxy, square rooms with no curtains or other softeners. All these things make audio systems **strain** to avoid picking up heavy **interferences**.

CONT.

- To enhance efficient multichannel communication, there is need to come up with a efficient, cost effective blind source system. This research aims to model one such system using neural networks.
- Artificial neural networks consist of numerous, simple processing units or “neurons” that we can globally program for computation. We can program or train neural networks to store, recognize and associatively retrieve patterns or database entries to solve combinatorial optimization problems, to filter noise from measurement data, to control ill-conditioned problem, in summary, to estimate sampled functions when we do not know the form of the functions.

BLIND SOURCE SYSTEMS

- **Non-linear mixtures**

Consider an $\boldsymbol{x}_i(t) = \{ x_1(t), x_2(t), \ldots x_n(t) \}^T$ **(1)**

The signal is non-linear memoryless mixture if it can be written as $\boldsymbol{x}_i(t) = \mathcal{T}\{a_{i1} * x_1(t) + a_{i2} * x_2(t) + a_{i3} * x_3(t), \ldots, + a_{mn} * x_n(t)\}$ where the mapping $\mathcal{T}$ consisting of the functions f_i is an unkown differentiable bijective mapping from one subset to another.

Taking away the time dependence variable the model will become;

$$\mathbf{x} = \mathcal{T}(\mathbf{s}) \qquad \textbf{(2)}$$

The question is: Is it possible to use signal processing to recover the sources s from the non-linear mixtures **(2)**?

CONT.

- **Results for Linear mixtures**

For linear mixtures (2) be comes;

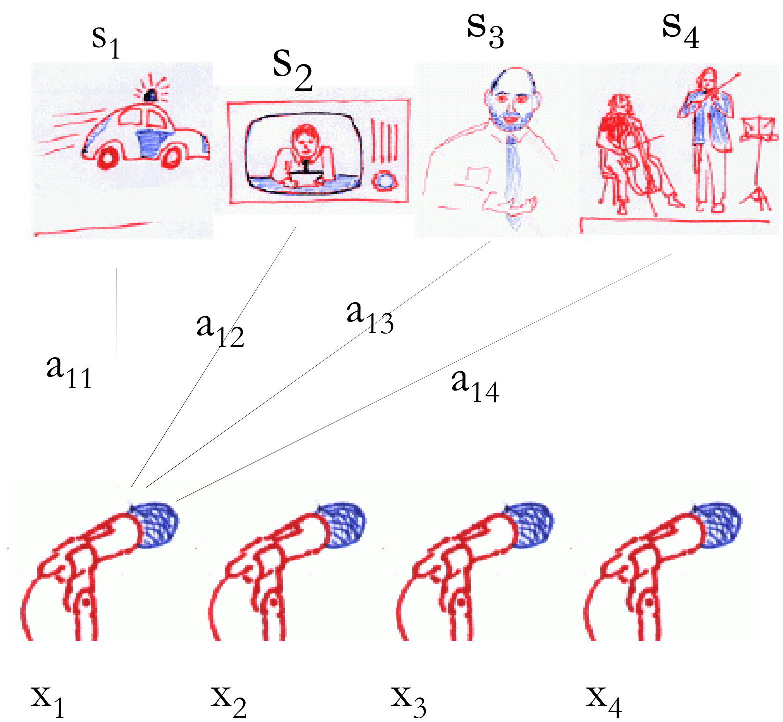
$$\mathbf{x} = \boldsymbol{A}\mathbf{s} \qquad \textbf{(3)}$$

where **A** is square non-linear mixing matrix. And therefore the source separation then of estimatating the non-singular matrix **A** is

$$\mathbf{y} = \mathbf{B(x)} = \mathbf{BA(s)} \qquad \textbf{(4)}$$

BLIND SOURCE SEPARATION BSS MODEL FOR SPEECH SIGNALS.

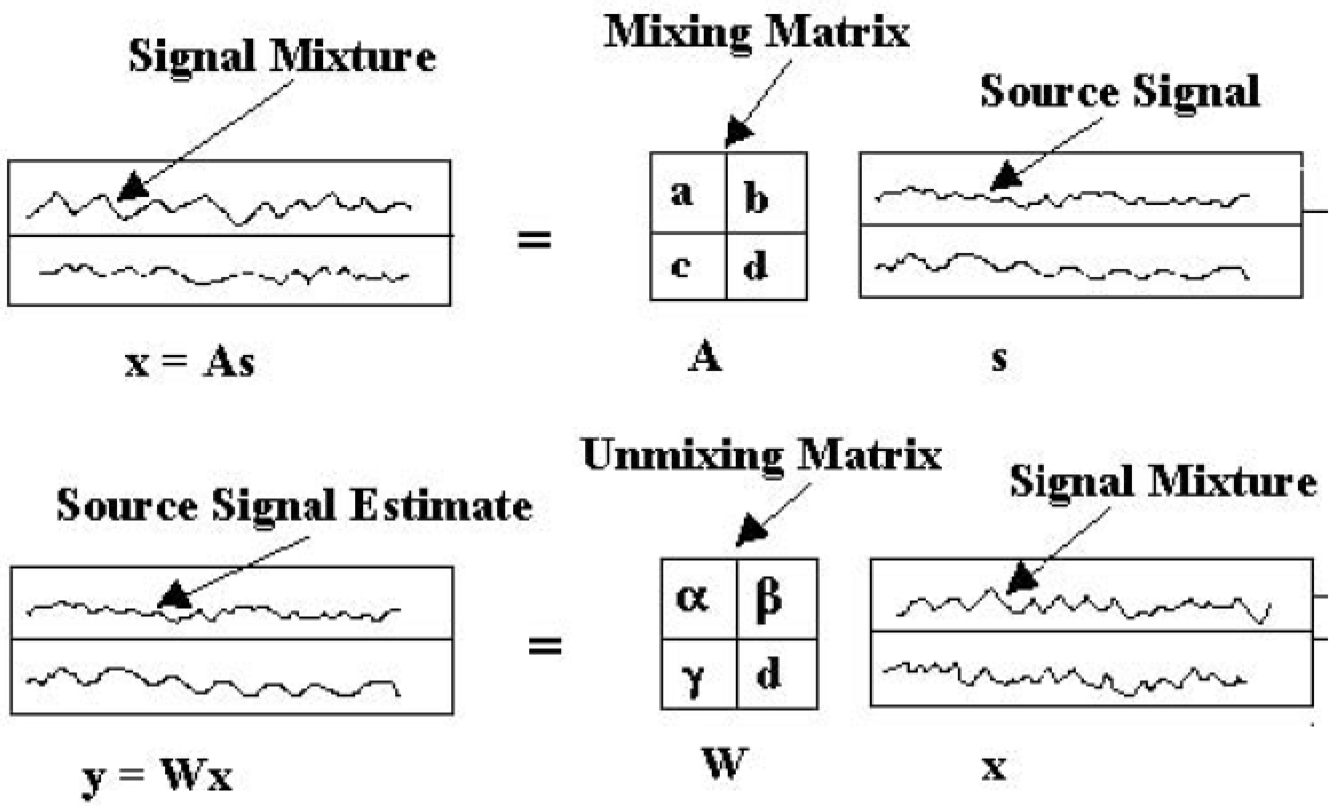
- Is a term used to describe the method of extrication (dis-connecting or dis-engaging)of underlying source signals from a set of observed signal mixture with little or no information as to the nature of those source signals.



INDEPENDENT COMPONENT ANALYSIS ICA MODEL

- $x_i(t) = a_{i1} * s_1(t) + a_{i2} * s_2(t) + a_{i3} * s_3(t) + a_{i4} * s_4(t)$
- Here, i=1:4.
- In vector-matrix notation, and dropping index t, this is

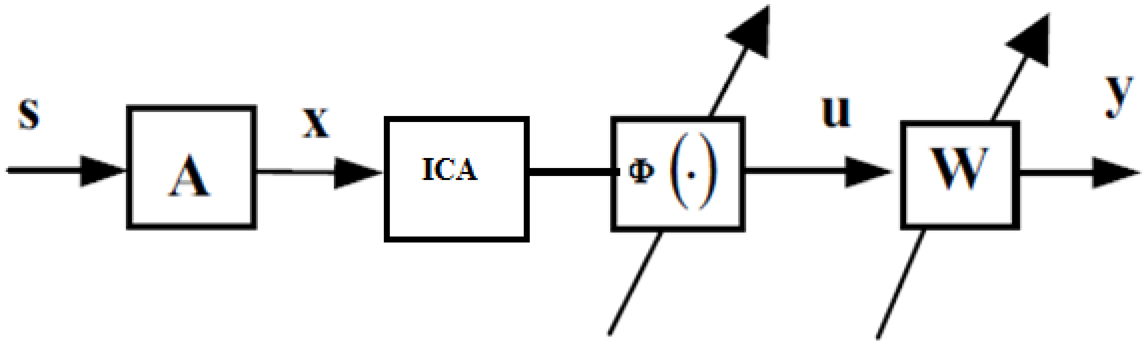
x = **A**



METHODS THAT HAVE BEEN FRONTED

- Matching pursuit- MP
- Principle Component Analysis- PCA
- Sparse Decompositions
- Independent Component Analysis- ICA

METHODOLOGY{THE MIXING AND DE-MIXING MODEL}



- The RBF neural network output layer is represented as, $y=\psi$ $(\mathbf{W\Phi})$.
- Where ψ (•) is a sigmoid function, W= (w_1, w_2..., w_n). By choosing a good on-line training algorithm having a good number of hidden computational units h,
- $y=\psi\,(\mathbf{W\Phi}) = \psi\,(\mathbf{W}\ \mathbf{\Phi}\ (\mathbf{x},\boldsymbol{\mu},\boldsymbol{\sigma}))=$ s.

ASSUMPTIONS TO BSS

- Statistically independent source signal
- High Kurtosis (High – order signal- 2^{nd} and 4^{th} orders)
- Ergodic (Stationary)…. Convolved(Differentiable between)

Therefore we will come with a signal processing technique to extract pertinent information from random signals using very little priori knowledge with aid of the above assumptions. This can be called signal estimation or filtering.

SIGNAL ESTIMATION

Estimation can be thought of a procedure made up of three primary parts:-

- Criterion Function
- Model
- Simulation of the proposed algorithm.

BSS - ICA

- Entropy is related to independence and since independence can not be measured but entropy can, we apply this tool in our method. Entropy of signal mixtures is constant and by mapping these signals maximizes this independence

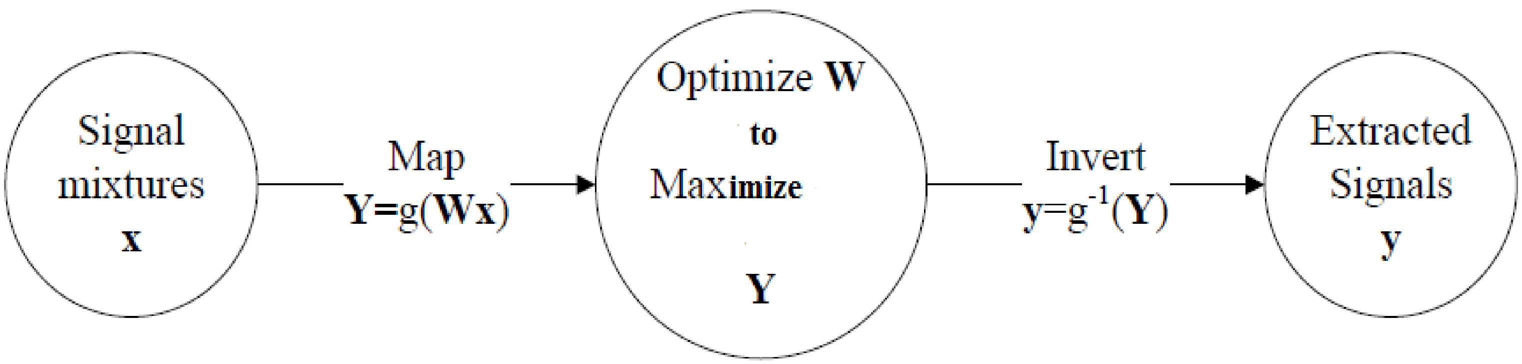
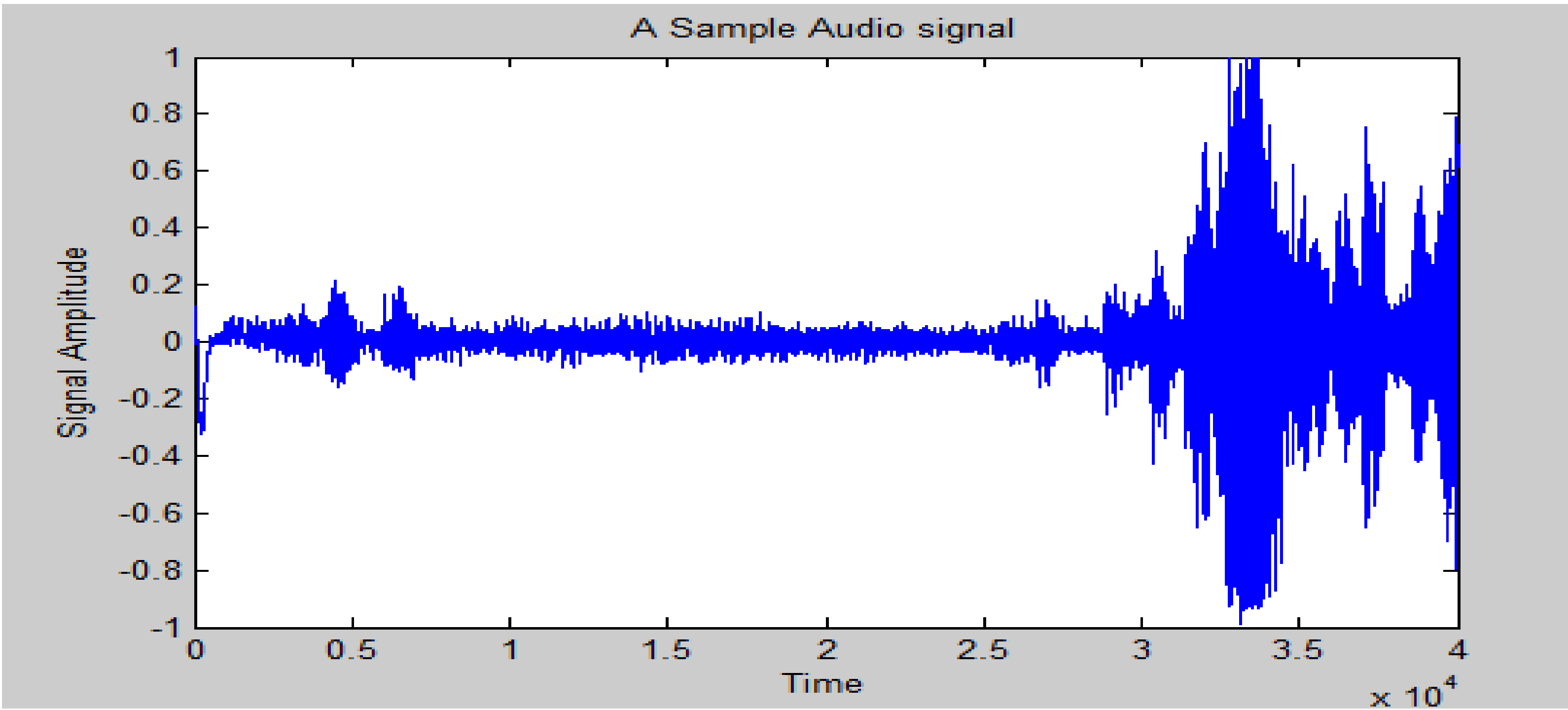
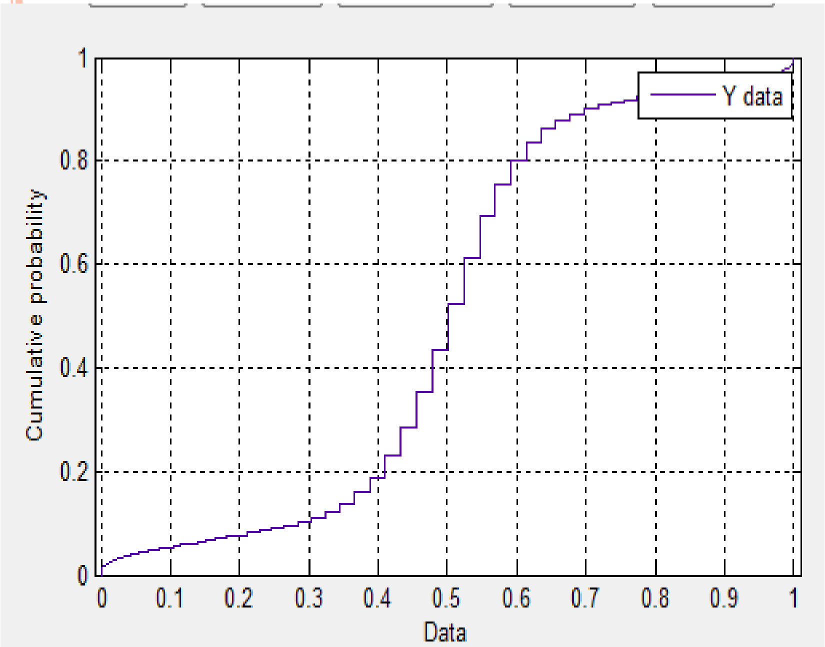


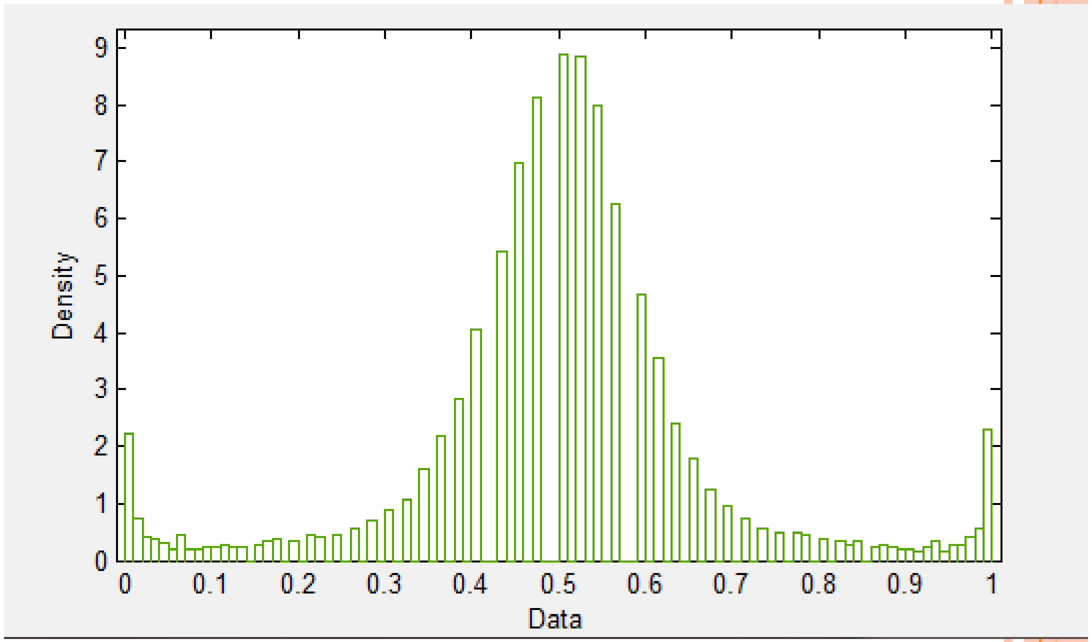
Figure BSS Strategy

ANALYSIS OF A SIMPLE AUDIO SIGNAL



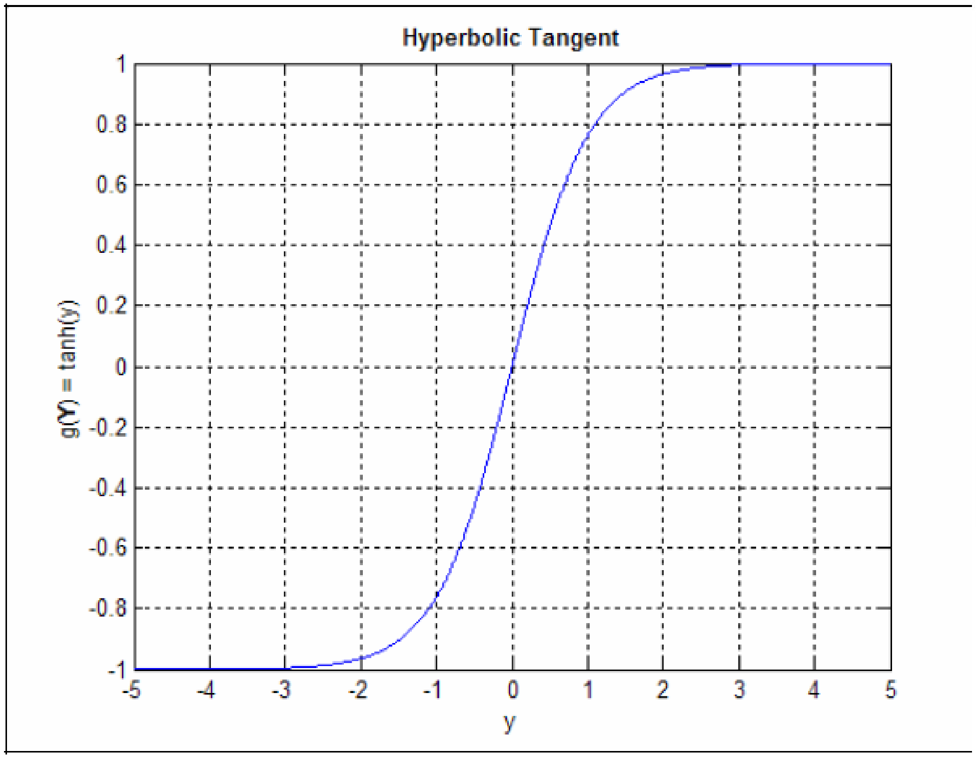
THE CDF AND PDF OF THE SAMPLE SIGNAL



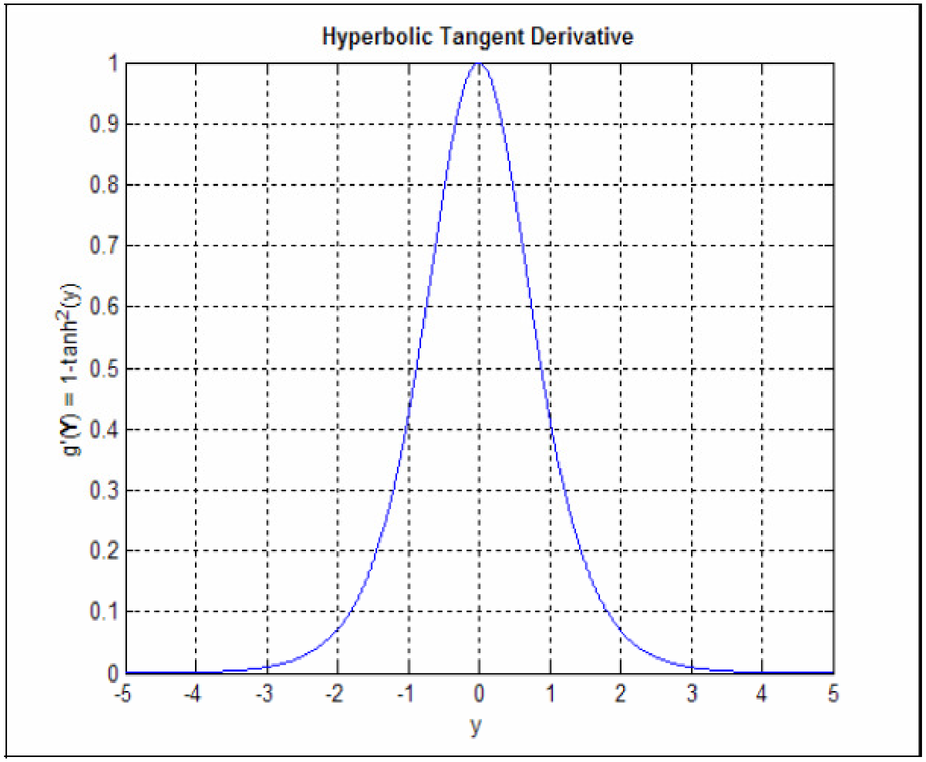


Non-linear functions for BSS

- The non-linear function should be judiciously selected to deal with the super-Gaussian, sub-Gaussian, stationary and non-stationary signals. The non-linear functions should be monotonic and invertible. The popular non-linearities used are logistic function and hyperbolic tangent function
- $Y = g(y) = \tanh(y)$



Hyperbolic tangent function.



Hyperbolic tangent derivative.

$$\mathbf{Y} = g(\mathbf{y}) = \tanh(\mathbf{y})$$

$$g'(\mathbf{y}) = \frac{d}{dy}\tanh(\mathbf{y}) = \operatorname{sech}^2(\mathbf{y}) = 1 - \tanh^2(\mathbf{y})$$

NONE-LINEAR MAPPING RELATIONS

$$P_Y(Z) = \frac{P_y(y)}{\frac{dZ}{dy}} = \frac{P_y(y)}{g'(y)} = \frac{P_y(y)}{P_s(y)},$$

$$P_Y(\mathbf{Z}) = \frac{P_y(\mathbf{y})}{\left|\frac{\partial \mathbf{Z}}{\partial \mathbf{y}}\right|},$$

$$P_y(\mathbf{y}) = \frac{P_x(\mathbf{x})}{\left|\frac{\partial \mathbf{y}}{\partial \mathbf{x}}\right|},$$

$$P_y(\mathbf{y}) = \frac{P_x(\mathbf{x})}{|\mathbf{W}|},$$

CALCULATING WEIGHTS BY MAXIMUM ENTROPY ME

Developed from Bell A. J and Sejnowski T. J **"Information maximization approach Algorithm"** called entropy to generate weights for RBF. The training of RBF with linear dependence on the output layer weights and non-linearity introduced in the input help to address the problem of local minima. The extracted signals ***y*** are obtained from signal mixtures ***x*** by optimizing the unmixing matrix **W**

The differential entropy of a signal is given by;

$$H(\mathbf{Y}) = H(\mathbf{x}) + \mathrm{E}\left[\sum_{i=1}^{N} \ln p_s(y_i)\right] + \ln|\mathbf{W}|,$$

$$h(\mathbf{Y}) = \mathrm{E}\left[\sum_{i=1}^{N} \ln g'(y_i)\right] + \ln|\mathbf{W}|,$$

CONT.

- This is the entropy associated with mapping of **x** to **Y.**
- Optimal matrix **W** is found using Gradient Descent on ***h*** by iteratively adjusting **W** in order to maximize the function ***h*.**

Gradient Entropy

By partial derivative of the entropy expression

Above; $$\frac{\partial h}{\partial \mathbf{W}_{ij}} = \mathrm{E}\left[\sum_{i=1}^{N} \frac{\partial \ln g'(y_i)}{\partial \mathbf{W}_{ij}}\right] + \frac{\partial \ln |\mathbf{W}|}{\partial \mathbf{W}_{ij}}.$$

Cont.

$$\nabla h = \mathbf{W}^{-T} + \mathrm{E}\left[\psi(\mathbf{y})\,\mathbf{x}^T\right],$$

- Thus the gradient descent rule, which in its most general form is

$$\mathbf{W}_{new} = \mathbf{W}_{old} + \eta \nabla h,$$

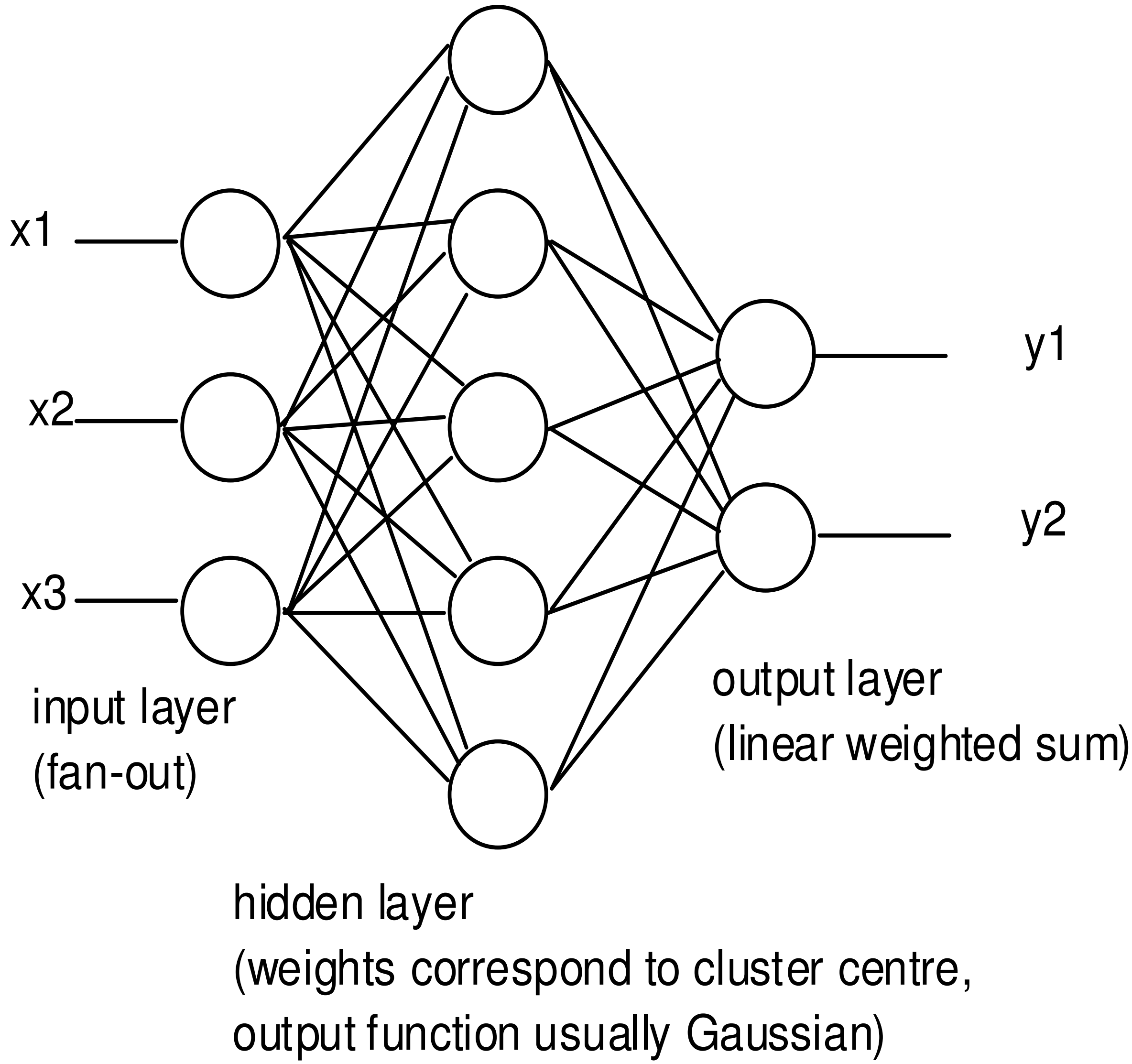
WHY MAXIMUM ENTROPY?

- While the assumption that the source signals are independent, independence of signals cannot be measured, entropy can.
- Entropy is related to independence in that maximum entropy implies independent signals. Therefore the objective of "Machine intelligent RBF" of mimicking the real world of a hybrid of linearity and non-linearity is easily attained by using the unmixing matrix **W** that is related to ME in the extracted signals, to train it

RBF NETWORK

- A function is approximated as a linear combination of radial basis functions (RBF). RBF's capture local behaviour of functions.
- Much of the inspiration for RBF networks has come from traditional statistical pattern classification techniques { Cover's Theorem}

CONT.


x1
x2
x3
y1
y2
input layer
(fan-out)
output layer
(linear weighted sum)
hidden layer
(weights correspond to cluster centre,
output function usually Gaussian)

CONT.

- For an n-output and n-output RBF network model. It consist three layers;

  input layer

  hidden layer

  output layer

- The neurons in hidden layer are of local response to its input and called RBF neurons while the neurons of the output layer only sum their inputs and are called linear neurons. The RBF network of Fig. above is often used to approximate an unknown continuous function which can be described by $\boldsymbol{\phi}: \boldsymbol{R}^n \dashrightarrow \boldsymbol{R}^n$ the mapping

  $$\mathbf{y}(\mathbf{x}) = \mathbf{W}\boldsymbol{K}(\mathbf{x}, \mathbf{p})$$

- where $\mathbf{W} = \mathbf{w}_{ij}$ is a $n \times M$ weight matrix of the output layer, $\boldsymbol{K}(\mathbf{x}, \mathbf{p})$ is a kernel function vector of the RBF network consisting of local functions.

CONT.

***K*(x, p)** can be represented as ***K*(*r*)** takes one of the sevearl forms such as

- Linear, cubic, thin plate spline, Gaussian, multiquatradic or inverse multiquatradic.
- For the proposed separating model we choose the conventional Gaussian kernel as activation of RBF neurons and it becomes:-

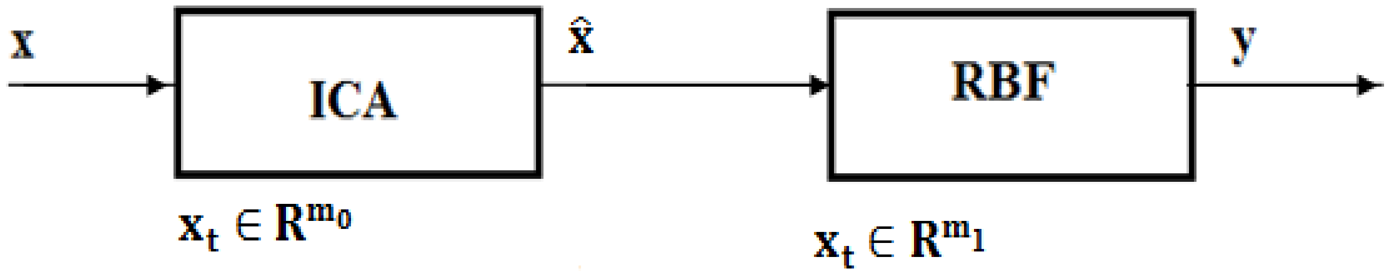
Gaussian Activation Kernel for RBF (non-Supervised and supervised)

$$K(x,p)=K(r)=\exp(-r^2/2)=\exp\{-\|x-\mu_i\|^2/2\sigma_i\}$$

$$K(x,p)=[1,\exp\{-\frac{\|x-\mu_1\|^2}{2\sigma_1}\},\ldots\ldots\exp\{-\frac{\|x-\mu_N\|^2}{2\sigma_N}\}]$$

P= $(\mu_1,\ldots\ \mu_N,\quad \sigma_1,\ldots\ \sigma_N)$ is the centre parameter and maximum distance for the Gaussian shape that makes it neither too narrow nor too flat

THE SCOPE OF WORK TO BE DONE.



- Designing a RBF with appropriate number of input, hidden and output units.
- Training RBF with the Weights related to ME as
- Check the performance level attained by this method as related to BSS, by using a good performance index esp after applying the following sample communication signal Polar NRZ
- Conclusion.

A SAMPLE COMMUNICATION SIGNAL MODEL: -(WAY FORWARD)

Once it is determined that RBF neural network analysis through Infomax Algorithm was fairly reliable for a number of signal, it is to be tested on a simple communication signal, the polar non-return to zero signal [11]. Digital baseband signals often use line codes to provide particular spectral characteristics of pulse train [15]. The most common codes used for mobile communication is polar *non-return-to-zero* NRZ because it offers simple synchronization capabilities. The modelling of the pdf which require a unit kronecker delta proofed difficult and we had to configure it as shown in Figures below.

Polar non-return to zero

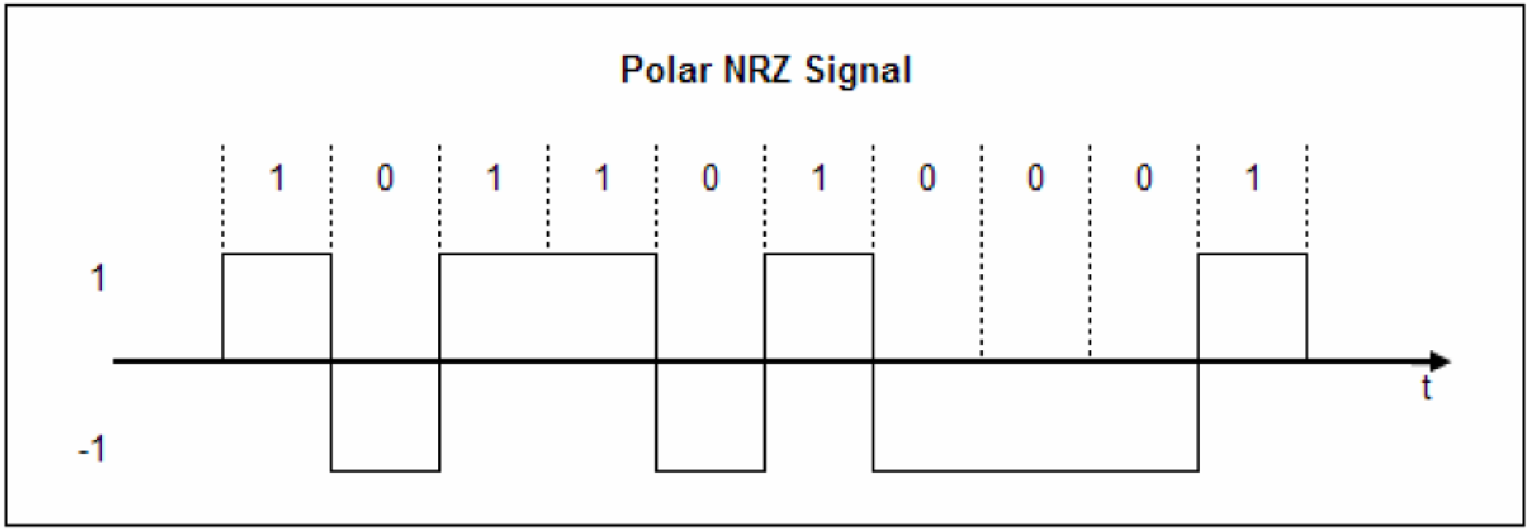
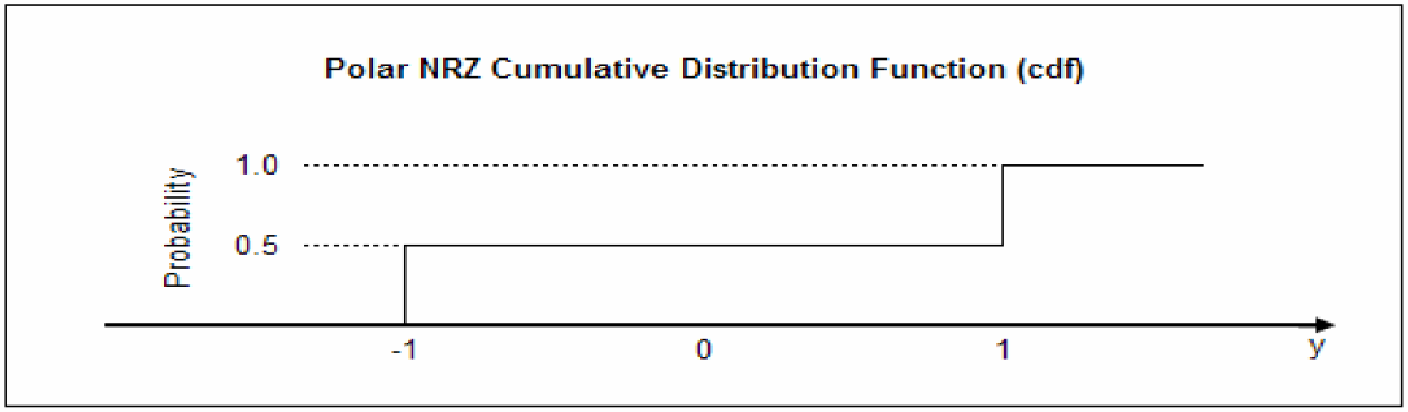
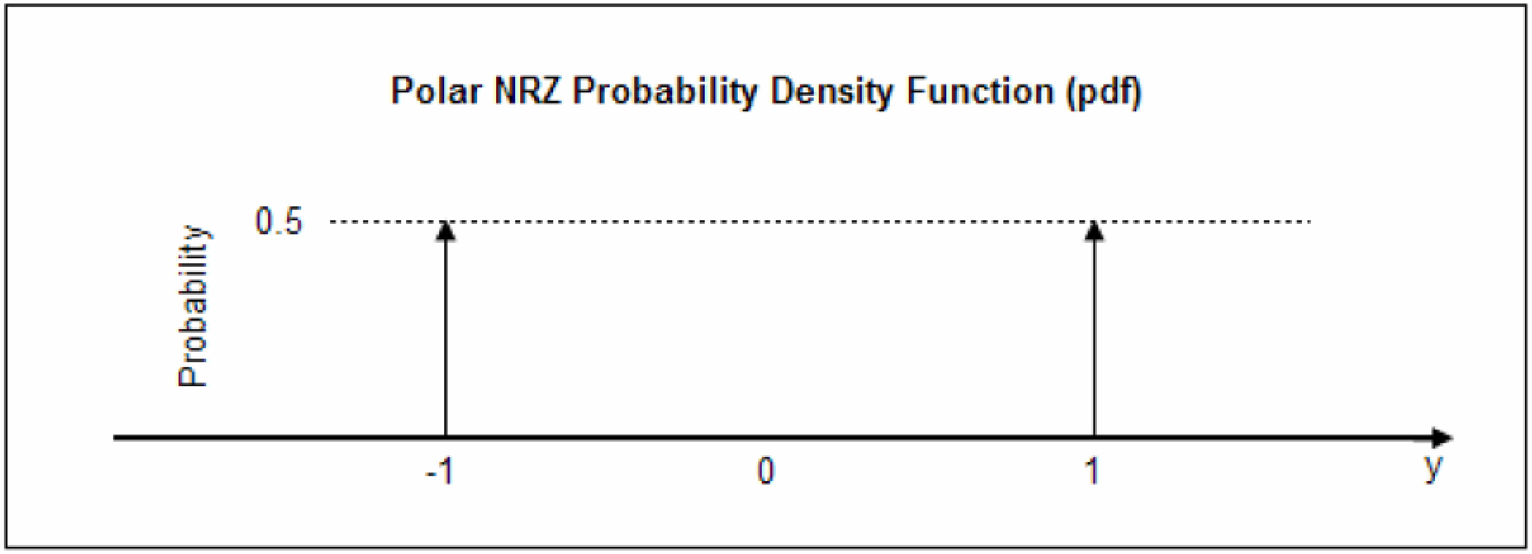


Figure Sample polar NRZ signal.

CDF AND PDF OF POLAR NRZ



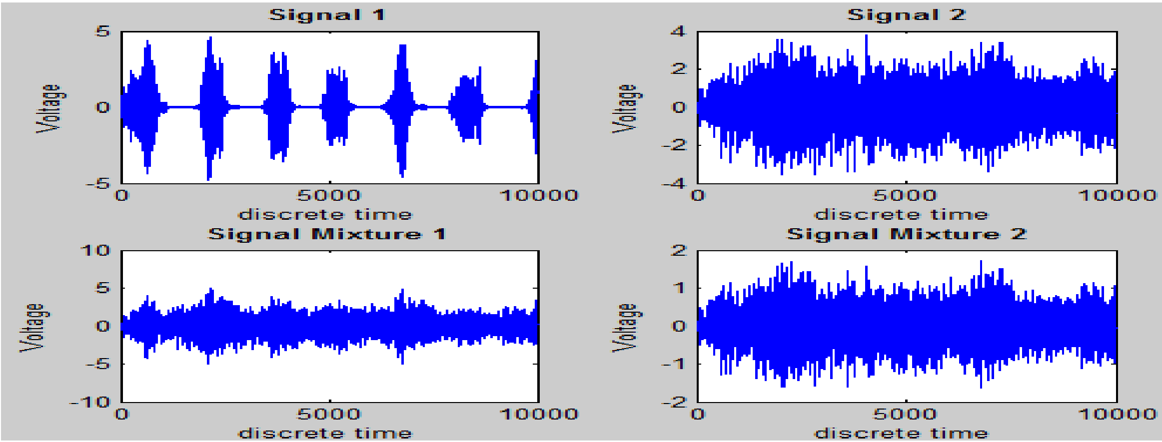
Theoretical polar NRZ cdf.

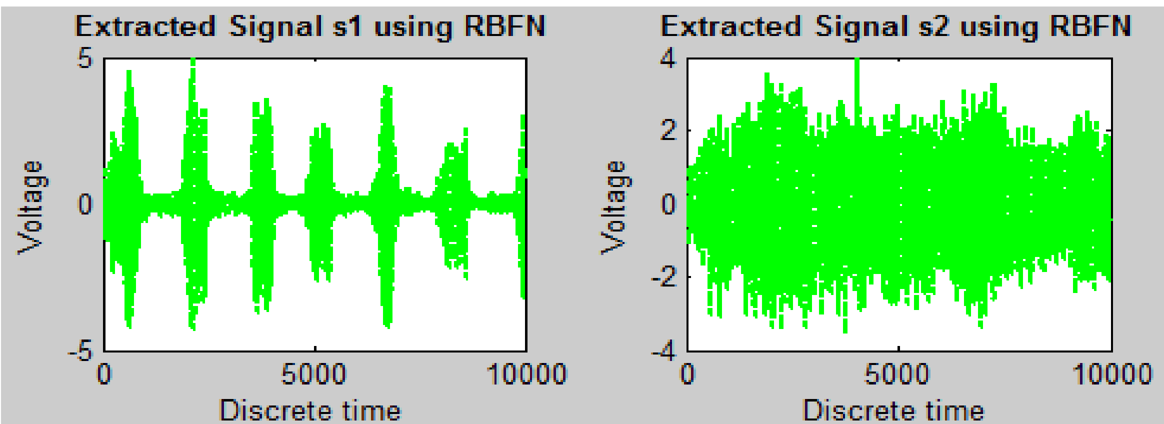


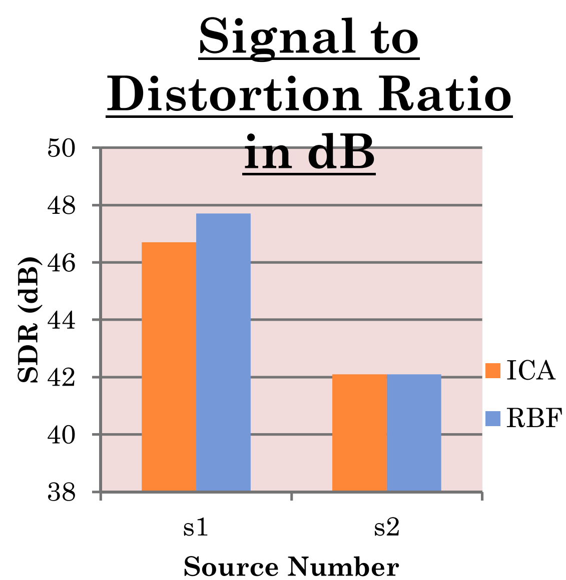
Theoretical polar NRZ pdf.

EXPECTED RESULTS

- Improved signal separation and strength.







TABULATED RESULST:

Table 1: Case 1 -Two Source Signal and Case 2 -Four Source Signals.

	SDR in (dB)		SIR in (dB)		SAR in (dB)		SDR in (dB)		SIR in (dB)	
	ICA	RBFN	ICA	RBFN	ICA	RBFN	ICA	RBFN	ICA	RBFN
S1	46.7	47.7	45.4	45.4	258.6	229.0	15.3	15.3	15.0	15.5
S2	42.1	42.1	41.0	41.0	244.8	232.6	4.8	4.9	4.8	4.9
S3					239.4	221.2	26.3	26.3	26.3	23.9
S4					242.3	227.6	6.9	6.9	6.8	7.0

TABULATED RESULTS CONTD.

Table 2: Case 3 Three NRZ Source Signal Estimation with Linear Mixture.

	SAR in dB		SDR in dB		SIR in dB	
Sources Signals	**ICA**	**RBFN**	**ICA**	**RBFN**	**ICA**	**RBFN**
S1	244.4	248.9	21.6	56.1	25.6	52.1
S2	247.3	239.6	41.3	34.5	44.8	37.5
S3	228.0	233.7	19.6	68.4	39.4	68.4

Table 3: Case 4: Three NRZ Source Signal Estimation with Non-linear Mixture.

	SAR in dB		SDR in dB		SIR in dB	
Sources Signals	**ICA**	**RBFN**	**ICA**	**RBFN**	**ICA**	**RBFN**
S1	234.1	233.1	4.1	6.7	4.1	6.7
S2	221.1	225.5	7.2	7.3	7.2	7.4
S3	232.1	218.8	28.9	57.4	25.7	57.4

THE END

THANK YOU

Indepedence of the Components

- Was slide 17

Problem Statement 1: 3 of 6

Problem

- Camera vision as an image analysis and interpretation problem.
- At low level is image pixel or feature labelling
- At high level is a contextual recognition and representation problem

Problem summary.

- Object identification and recognition
- Depth from 2D - Visual Ranging
- Spartial localization - Visual Tracking
- Processing Latencies - Computational speed
- Automated Adaptability to different conditions

Goals and Objectives I: 4 of 6

Main Objective

- To overcome critical limitations inherent in monocular vision.

Specific Objectives

- To Develop algorithm for visual tracking and ranging of road and objects employing virtual cameras, correlation techniques, Markov random fields and adaptive template updating mechanism.

- Optimize the adaptability and processing speed through modularization, parallelization and multithreading algorithm harnessing the power and parallel structure of neural networks and multi-core processors.

Justification of the study I: 5 of 6

- Little have been done on Monocular vision
- Stereo vision to date put less emphasis on real-time processing
- Promise of high computational speeds
- Promise of a more robust and adaptive vision system
- Application on a driver assistance system (DAS).
- Application on other camera vision or surveillance systems

Scope I: 6 of 6

- Algorithm development and deployment in C code targeting FPGA
- Road features recognition and representation
- Front end vision.
- Straight road sections and long curved sections

Visual Tracking II: 1 of 6

- Differential Digital Image Trackers
- Particle filter or condensation based tracker
- Active contours tracker
- Mean shift tracker
- Integral histogram based tracker
- Covariance tracker
- Edge tracker
- Centroid tracker
- Corr t
  Corr

Correlation

$$\rho(x_i, x_{i'}) = \frac{\sum_j (x_{ij} - \bar{x}_i)(x_{i'j} - \bar{x}_{i'})}{\sqrt{\sum_j (x_{ij} - \bar{x}_i)^2 \sum_j (x_{i'j} - \bar{x}_{i'})^2}}$$

Visual Ranging II: 2 of 6

Applications

- Saxena et al. Estimated 1-D distances for driving a remote control car autonomously
- Nagai et al. performed Surface reconstruction from single images
- Gini & Marchi used single-camera vision to drive an indoor robot

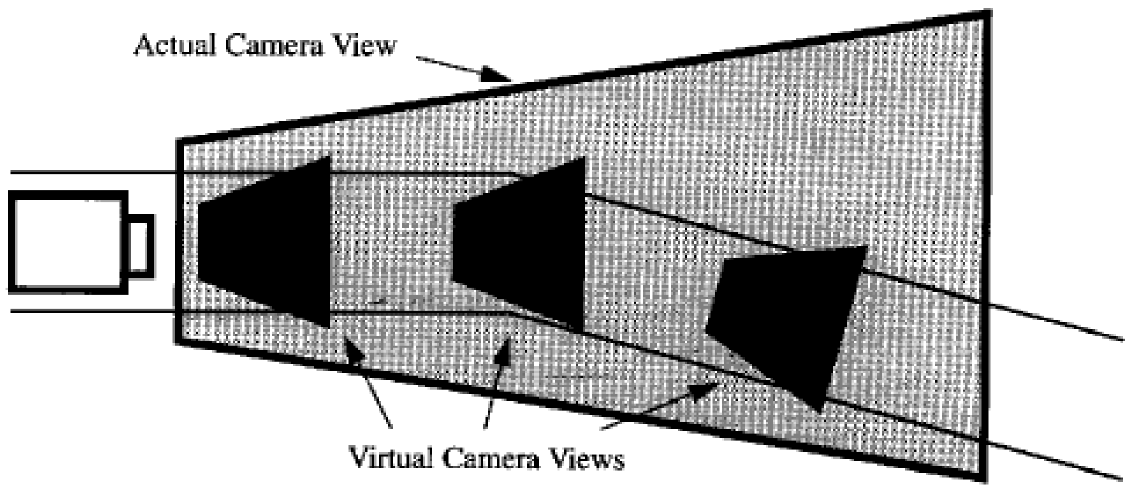
Methods

- Range from stereo
- Range from vergence
- Range from focus/ Defocus
- Affine structures
- Shape from shading
- Supervised training by Markov Random Fields(MRF)
- Neural Net MRF/CRF method utilizing Perspective concept and Geometrical transformation

VIRTUAL CAMERAS II: 3 OF 6

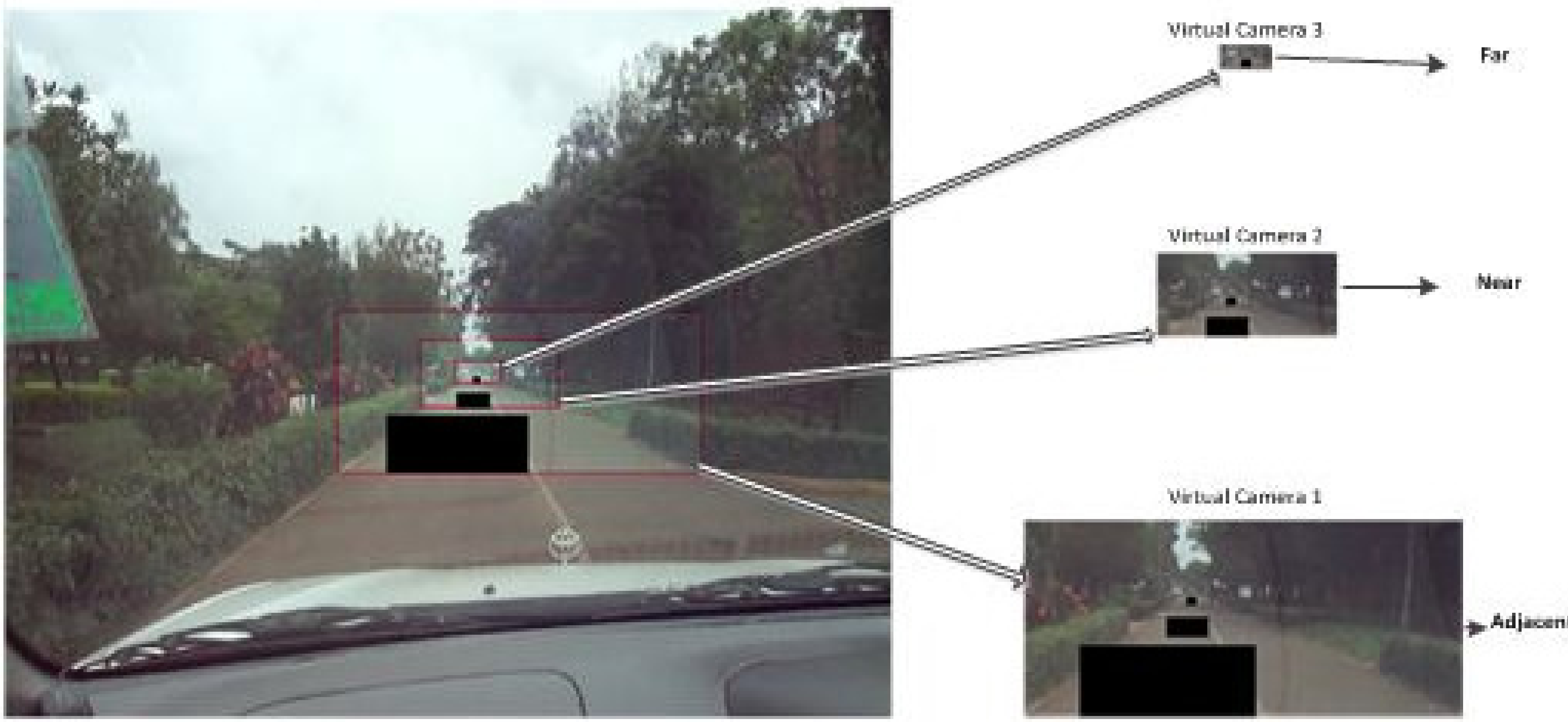
- Birds Eye-View (Flat world Model)
  (Wallace R., 1986) (Lotufo R., 1990)

- Virtual Cameras
  (Todd, 1994)



Vertical section Virtual camera views

Virtual camera Concept Tracking
the Road, depth and the Objects



- Correlation techniques is used in Road and objects segmentation.
- Markov Random Fields used in initialization of the virtual

Computational Speed (Latencies) II: 5 of 6

- Kernel based algorithms(Kernel RBN and SVM)

  (Vapnik, 1995)
- Pulse-coupled Neural networks

  (Kinser and Lindblad, 2005)
- Multiple classifier algorithms

  (Many literals)
- Binary Neural Networks

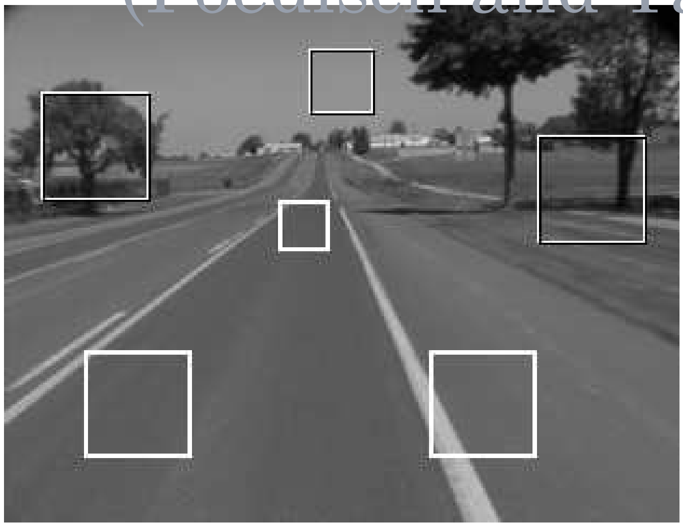
  (Austin 1996)
- Sub-sampling and distributed computing

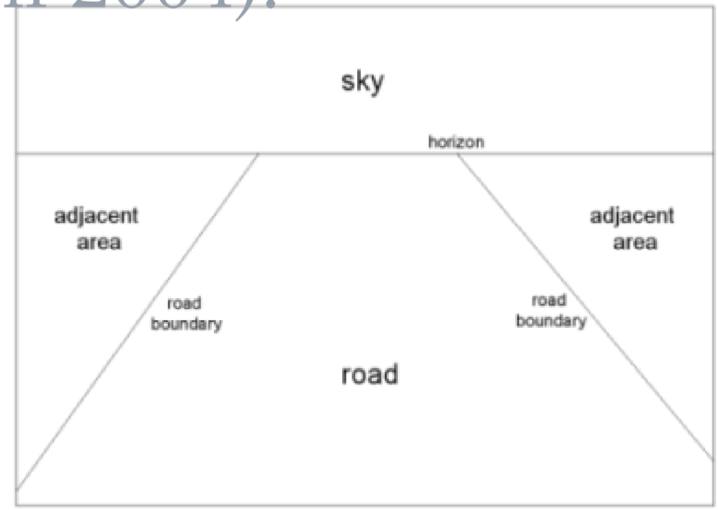
  (ALVINN Project)
- Modular Neural networks

  (Evangelia and Cooley, 2000) (Schmidt,1996)

Automated Adaptability II: 6 of 6

- Adaptive Road Detection Using Neural Networks
  (Foedisch and Takeuchi 2004):





- Discriminated learning of the road features only
  {Correlation Method}

Methodology III: 1 of 12

- Data Collection
- Data Analysis
- Algorithm development and evaluation
- Algorithm Optimization and evaluation
- Deployment in Fixed point C code & Evaluations

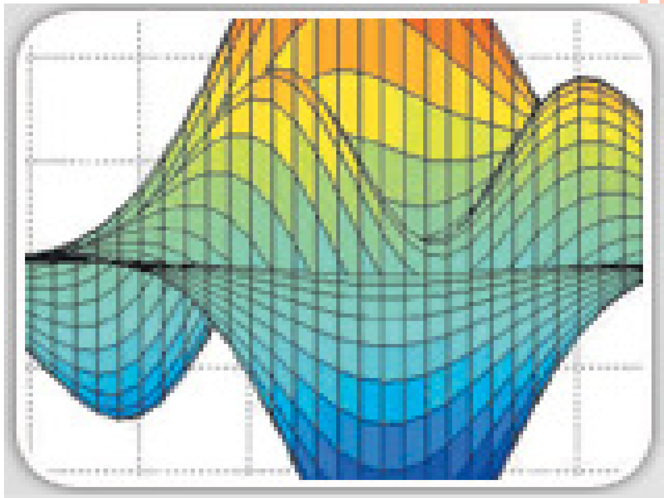
DATA COLLECTION III: 2 OF 12

- Real Road Driving video recording

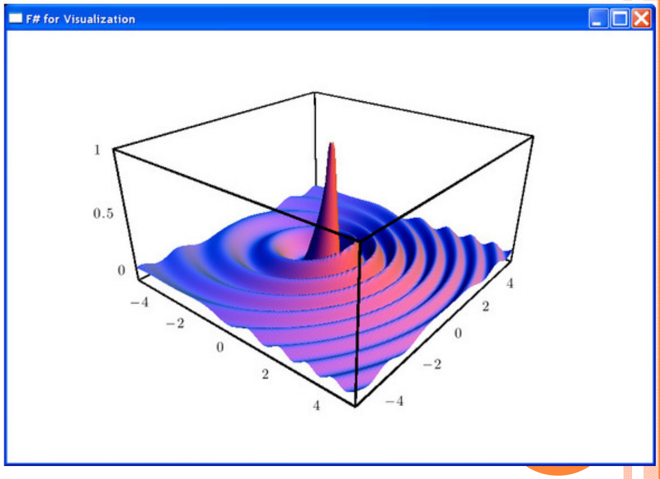
- Video frames binned along Vehicle speed

Data Analysis (Matlab & F#) III: 3 of 12

- Color Histogram indexing method
- Wavelets methods
- Watershed methods
- K-Means clustering methods
- Morphological processing
- Filters (Noise reduction, Edge enhancement)
- Hough transforms



Matlab Interactive Kit

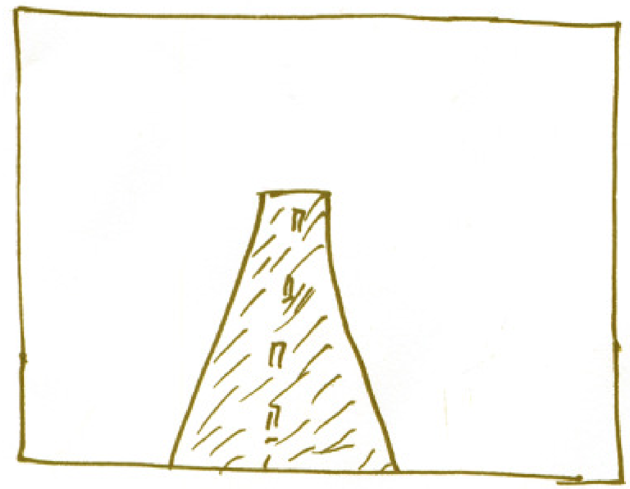


F# For Visualization

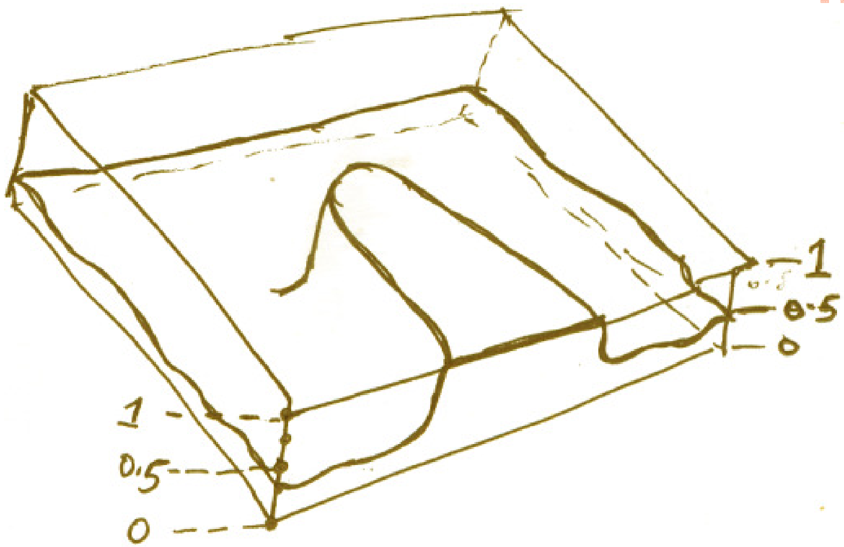
Road Segmentation Algorithm Dev. III: 4 of 12

Module1

- Image preprocessing algorithms
- Edge enhancing algorithms
- Correlation algorithm
- Template updating algorithm



2D Road image

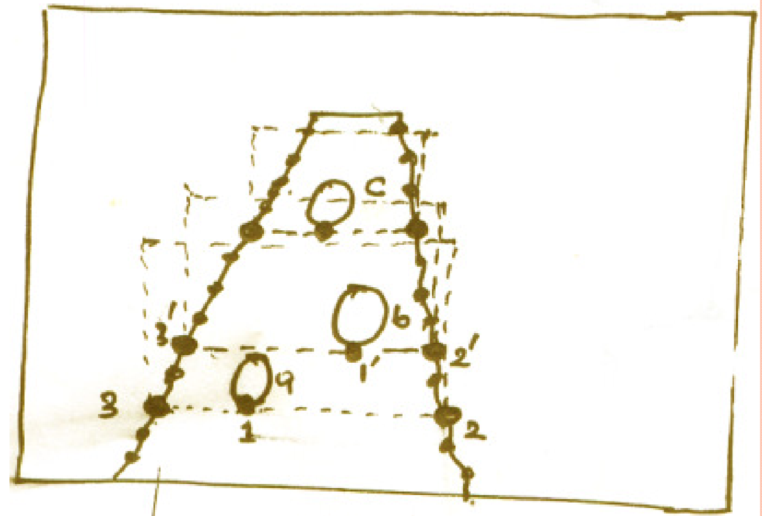


Correlation Surface

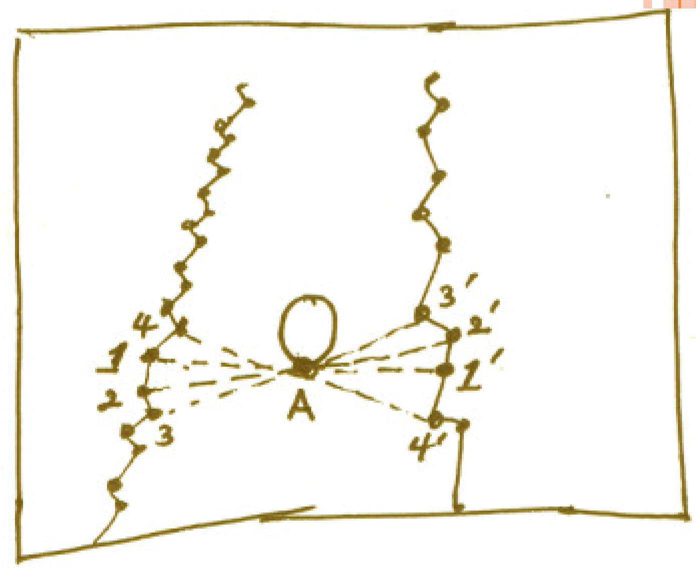
Virtual Cameras Algorithm Dev. III: 5 of 12

Module 2

- Objects lower edge node identification
- Virtual cameras window initialization
- Windows Normalization
- Ranging algorithm



Virtual cameras initializat

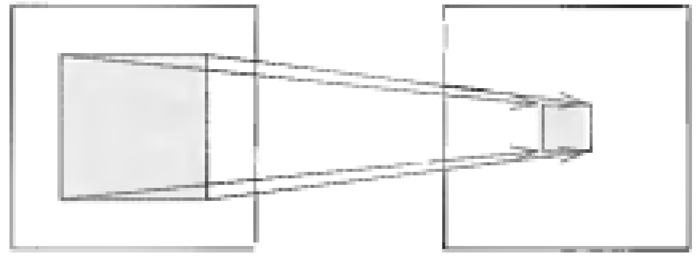


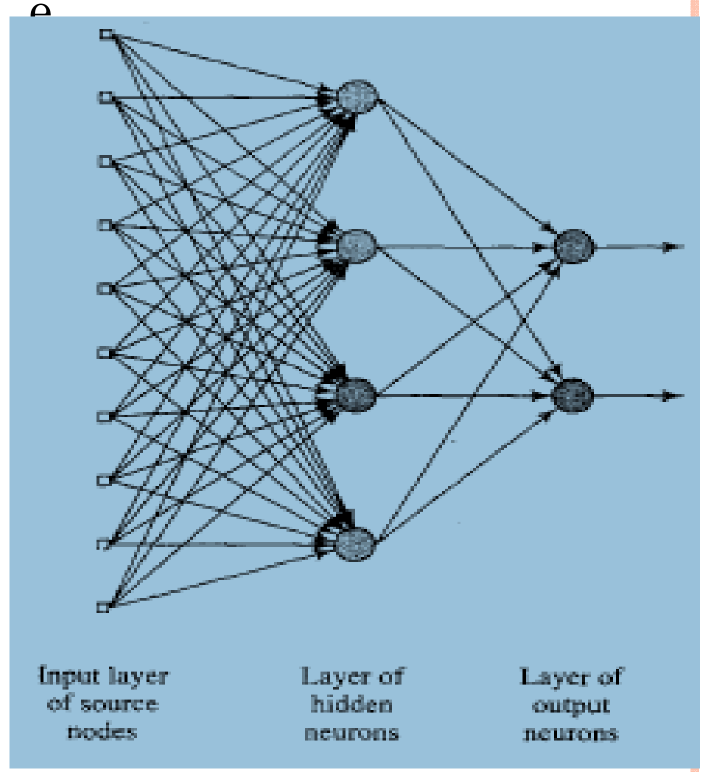
Selecting Best Base line using

Recognition Algorithm Dev. III: 6 of 12

Module 3

- Recognition algorithm
- Output Representation
- Feed Forward Neural network
- Hidden Markov Models

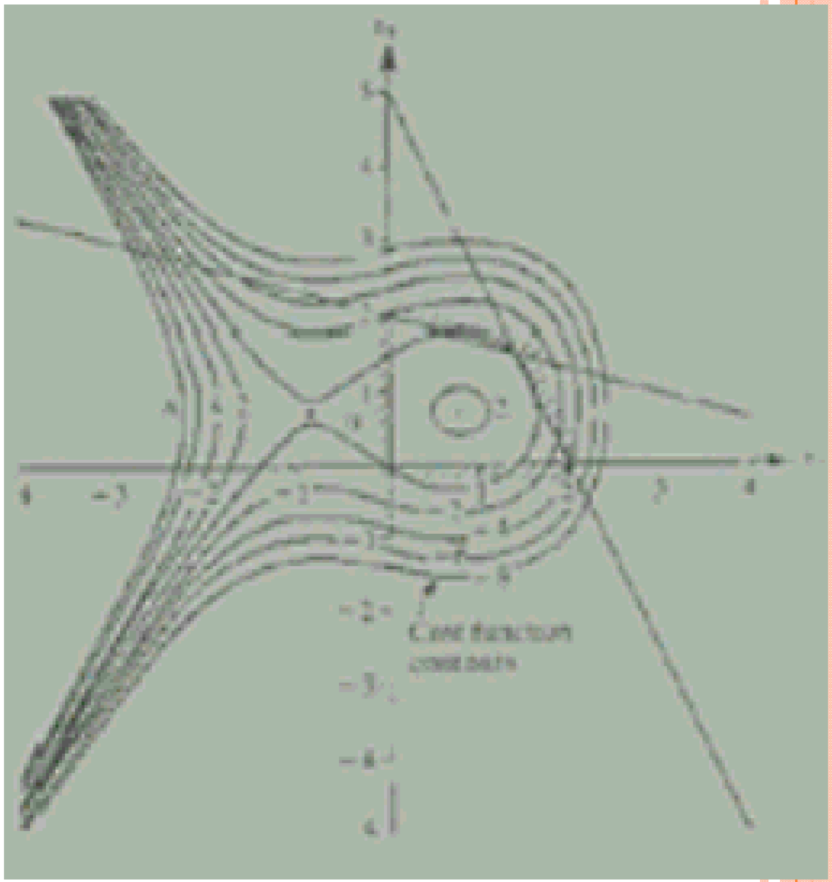




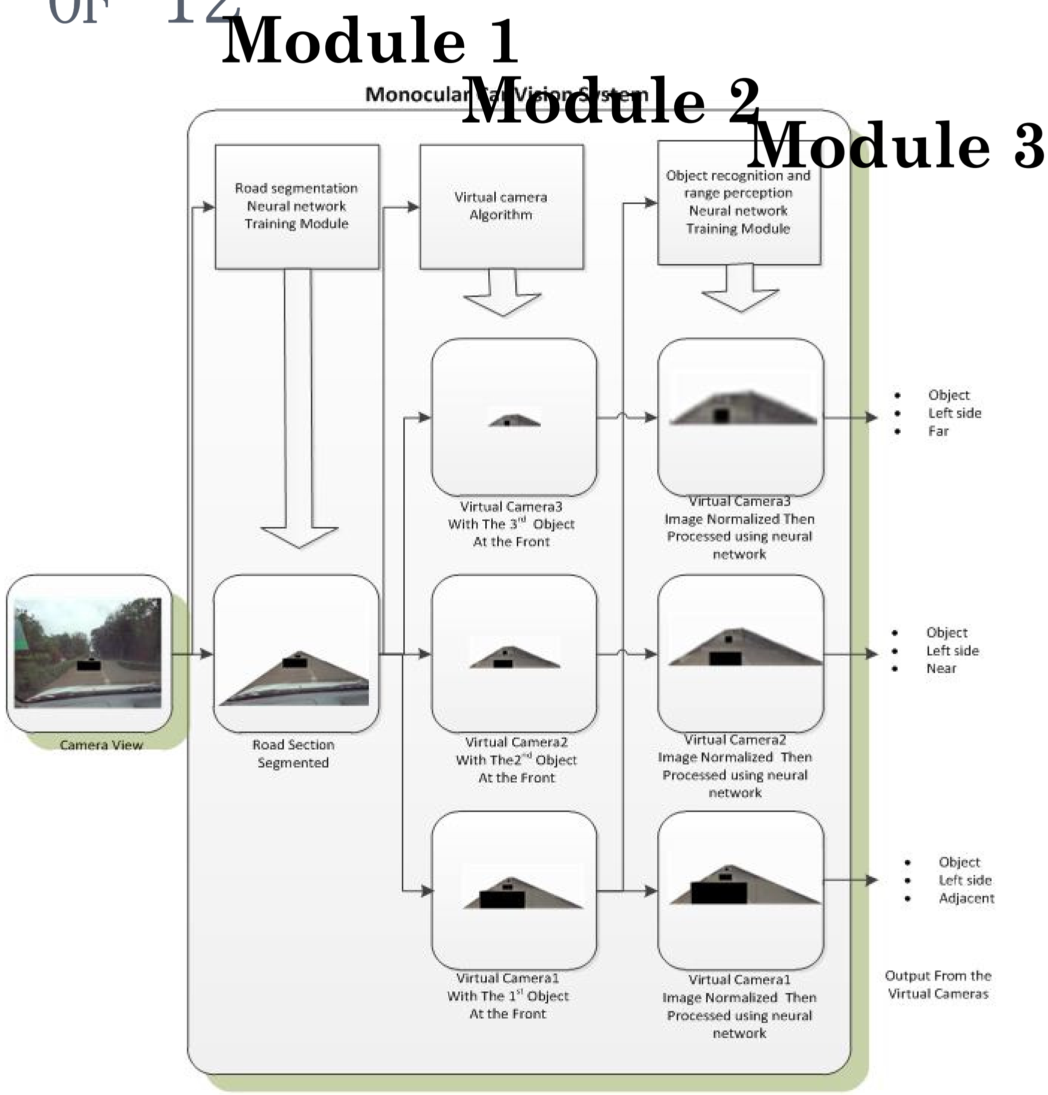
Optimization & Deployment III: 7 of 12

Optimization

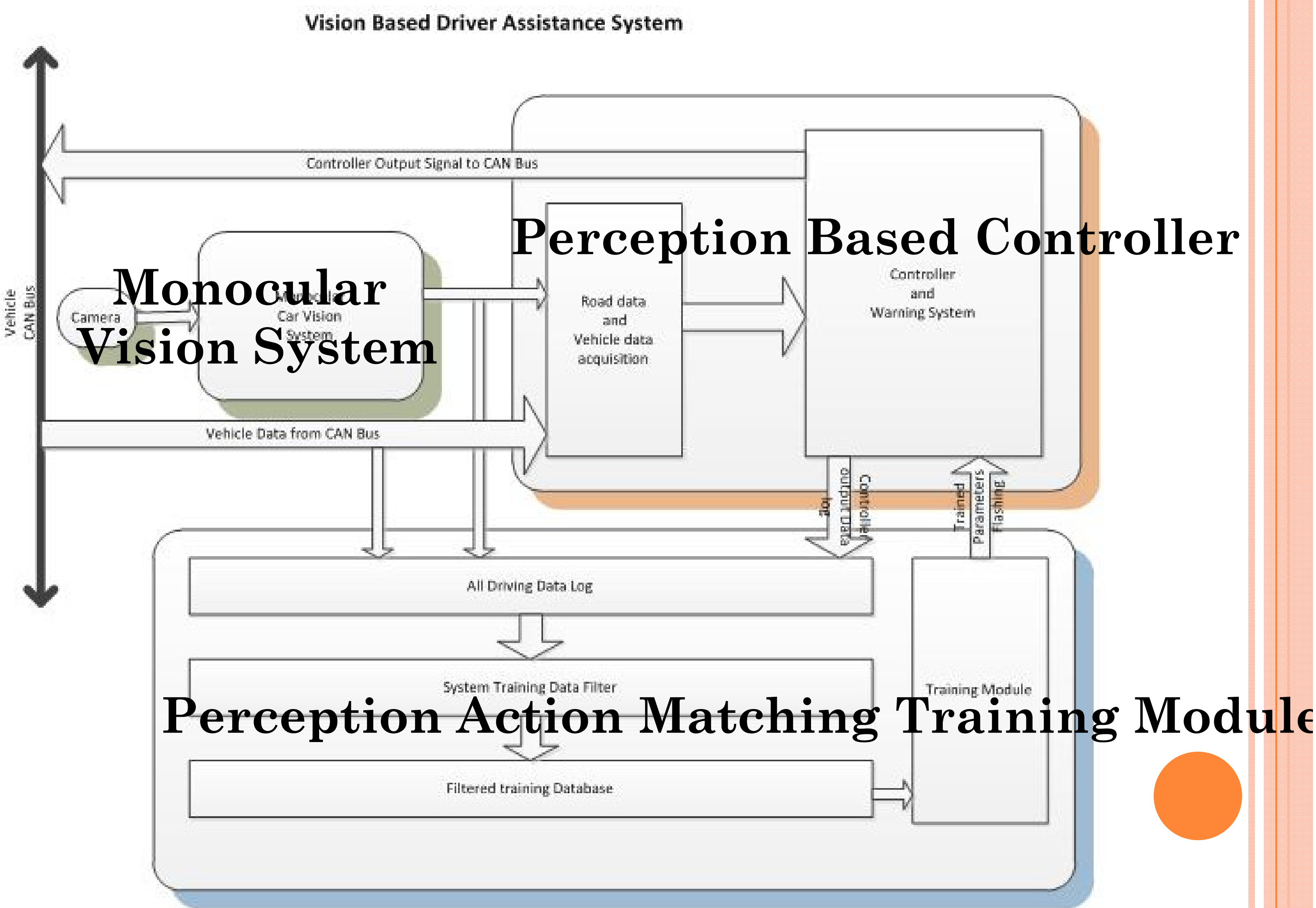
- Modularization
- Parallelization
- Vectorization
- Multithreading or Pipelining
- Fixed Point C code deployment and evaluations
- Every module will be evaluated on Speed and Accuracy (Errors)



PROPOSED DESIGN III: 8 OF 12



TARGET APPLICATION II



CONTRIBUTIONS OF THE STUDY III: 10 OF 12

- Vertical virtual cameras concept
- Ranging in 2D images
- Computational speed acceleration
- Neural Net MRF Models
- Neural Net Correlation Function
- Road sensor for car vision

Requirements

III: 11 of 12

No.	Particulars	Estimated Cost (Kshs.)
1	DCAM camera Hire	10,000
2	Car Hire and Fuelling	10,000
2	Documentation, printing, photocopies & any other stationery	15,000
3	Text Books	25,000
4	Electronic journals and internet	15,000
Total cost		75,000

WORK SCHEDULE III: 12 OF 12

Activity	Aug 2010	Sep 2010	Oct 2010	Nov 2010	Dec 2010	Jan 2011	Feb 2011	Mar 2011	Apr 2011
Literature review									
Proposal conception and drafting of the proposal									
Proposal defending			Seminar	Seminar	Defense				
Design simulation									
Thesis writing									