Digital processing of biomedical signals and images

- Scientific bases and application examples-

Prof. Nizamettin AYDIN, PhD

naydin@yildiz.edu.tr

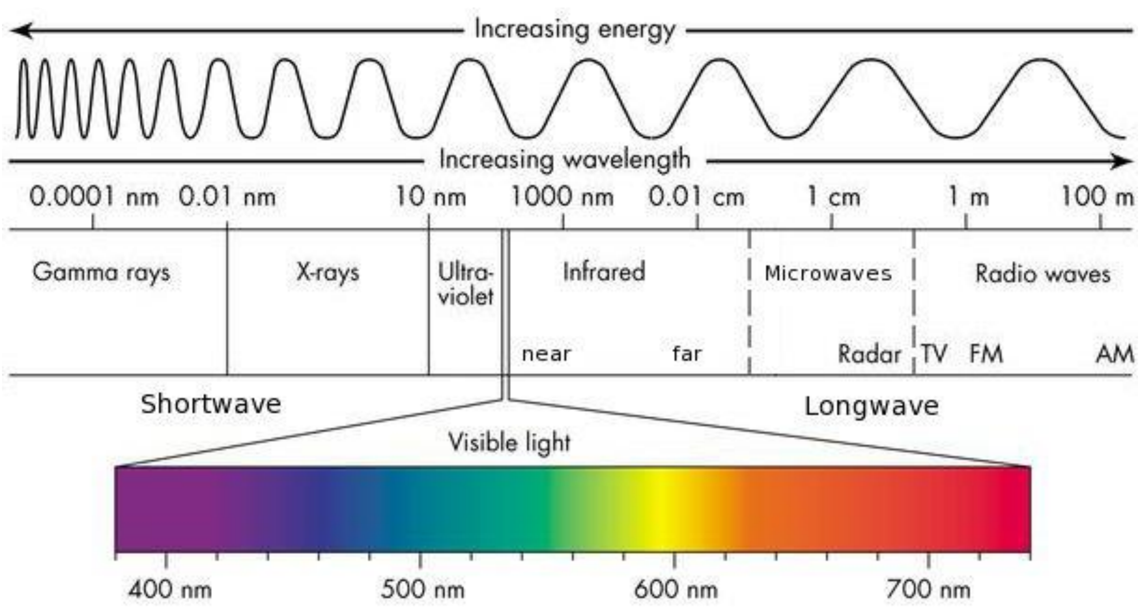
http://www.yildiz.edu.tr/~naydin

Signal, Image, Video

- A signal
  - a pattern of variations of a physical quantity that can be manipulated, stored, or transmitted by physical process.
  - an information carrying variable
  - one dimentional funtion $f(t)$
- An image
  - two-dimensional function $f(x,y)$, where x and y are the spatial coordinates,
  - the amplitude of f at any pair of coordinates (x,y) is called the intensity of the image at that level.
- A video
  - three-dimensional function $f(x,y,t)$, where x and y are the spatial coordinates, and t is the time.

The principal sources

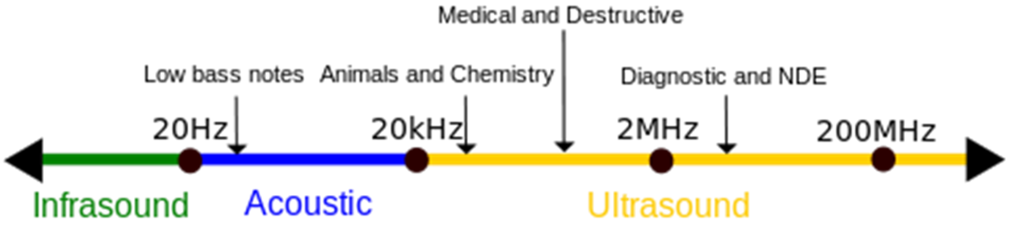
- The principal source for the signal is any measurable quantity
  - Electrical, physical, chemical, biological, etc.
- The principal source for the images is the electromagnetic (EM) energy spectrum.



- Gamma rays
- X- rays
- Ultra-violet
- Visible light
- Infrared
- Microwaves
- Radio Waves

The principal sources

- Ultrasound Imaging
  - Ultrasonic spectrum

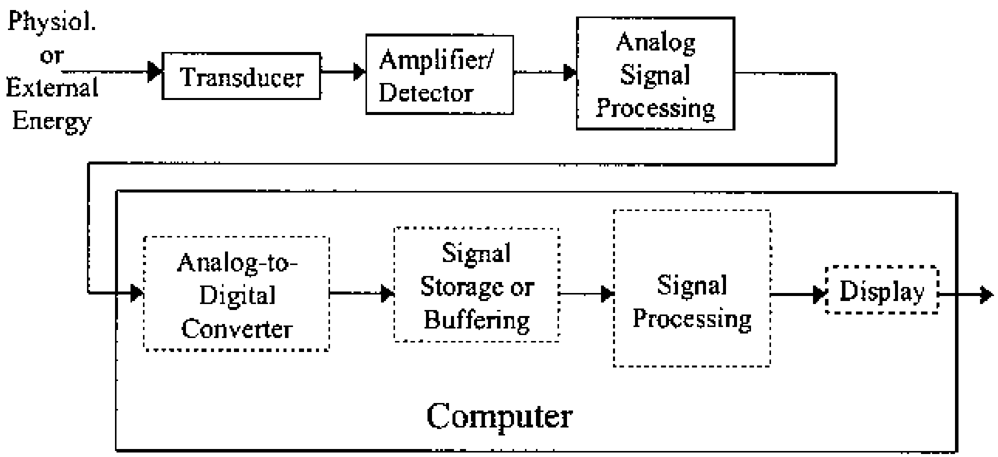


  - Sound (Acoustic)
    - Mechanical radiant energy that is transmitted by longitudinal pressure waves in a material medium and is the objective cause of hearing
  - Ultrasound
    - Vibrations of the same physical nature as sound but with frequencies above the range of human hearing
  - Infrasound
    - Vibrations of the same physical nature as sound but with frequencies below the range of human hearing

	Lowest		Highest
Turtle	20	-	1.000
Goldfish	100	-	2.000
Frog	100	-	3.000
Pigeon	200	-	10.000
Sparrow	250	-	12.000
Human	20	-	20.000
Chinpanzee	100	-	20.000
Rabbit	300	-	45.000
Dog	50	-	46.000
Cat	30	-	50.000
Guinea Pig	150	-	50.000
Rat	1.000	-	60.000
Mouse	1.000	-	100.000
Bat	3.000	-	120.000
Dolphin	1.000	-	130.000

Signal Processing

- involves the analysis of information captured through instruments that measure a 1D variable to provide useful information upon which a decision can be made.
  - Engineers are discovering new ways to process these signals using a variety of mathematical formulae and algorithms.
- A typical signal processing system



  - a great deal of activities takes place around the *signal processing block* in this system
  - However, advances in peripheral systems such as sensor and display technologies are also important for overall system improvement

Image Procesing Levels

- low-level processes

  Input: Image, Output: Image

  - Examples: Noise removal, image sharpening

- mid-level processes

  Input: Image, Output: Attributes

  - Examples: Object recognition, segmentation

- high-level processes

  Input: Attributes, Output: Understanding

  - Examples: Scene understanding, autonomous navigation

Biomedical signal processing

- involves the analysis of information captured through physiological instruments that measure
  - heart rate, blood pressure, oxygen saturation levels, blood glucose, nerve conduction, brain activity, etc.

  to provide useful information upon which clinicians can make decisions.
- By using more sophisticated means to analyze what our bodies are saying, we can potentially determine the state of a patient's health through more noninvasive measures.

Biomedical signal processing

- Biomedical Signal Processing combines many science and engineering disciplines such as
  - mathematics
  - physics
  - electrical engineering
  - computer engineering
- An advance in these fields also contributes to the advances in biomedical signal processing.

Biomedical Signal Processing (past)

- In the past, processing of biomedical signals meant mainly
  - filtering of signals for removing noise and power lines interference,
  - spectral analysis to understand the frequency characteristics of signals,
  - modeling for feature representation and parameterization.

Biomedical Signal Processing

- Physicians can make decisions by examining a patient.
- However, acquisition and processing of biomedical signals has become more and more important to the physician.
- The main reasons for this development
  - the growing complexity of the biomedical examinations,
  - the increasing necessity of comprehensive documentation,
  - the need for automation in order to reduce costs.

Biomedical Signal Processing (current)

- Recent trends have been toward quantitative or objective analysis of physiological systems via signal analysis and AI.
  - Analysis of signals accomplished by humans has many limitations
- Computer analysis of these signals could provide objective strength to diagnoses.
- Different techniques can be used to analyze a biomedical signal
  - filtering, adaptive noise cancellation, time-frequency analysis, pattern recognition, machine learning, etc.

Recent Technological Advances

- Recent Technological Advances
  - Sensor advances in medicine and biology
  - Nanotechnology
  - Advanced implants
  - Solar and light powered devices
  - MEMS
  - Wireless technologies
  - AI

  enables us to acquire more complex data
- Results in huge amount of data to be processed
  - Data science

Examples-1

- Advances in microarray technologies results in vast amount of data to be processed.
- Two approaches :
  - Computational
    - Using new computing systems, such as
      - Reconfigurable computing system / FPGA
      - GPU
  - Algorithmic
    - Applying well known signal processing techniques to a new data type
    - Applying AI and Data Science concepts

- H. Hussain, K. Benkrid, A. Ebrahim, A.T. Erdogan, and H. Seker, "Novel Dynamic Partial Reconfiguration Implementation of K-means Clustering on FPGAs: Comparative Results with GPPs and GPUs", Volume 2012, *International Journal of Reconfigurable Computing*, 2012.
- Faella, John, "On Performance of GPU and DSP Architectures for Computationally Intensive Applications" (2013). Open Access Master's Theses. Paper 2. http://digitalcommons.uri.edu/theses/2
- Chrysostomou, C., Seker, H., Aydin, N., Haris, P.I., "Complex Resonant Recognition Model in analysing Influenza a virus subtype protein sequences," Information Technology and Applications in Biomedicine (ITAB), 2010 10th IEEE International Conference on , vol., no., pp.1-4, 3-5 Nov. 2010.

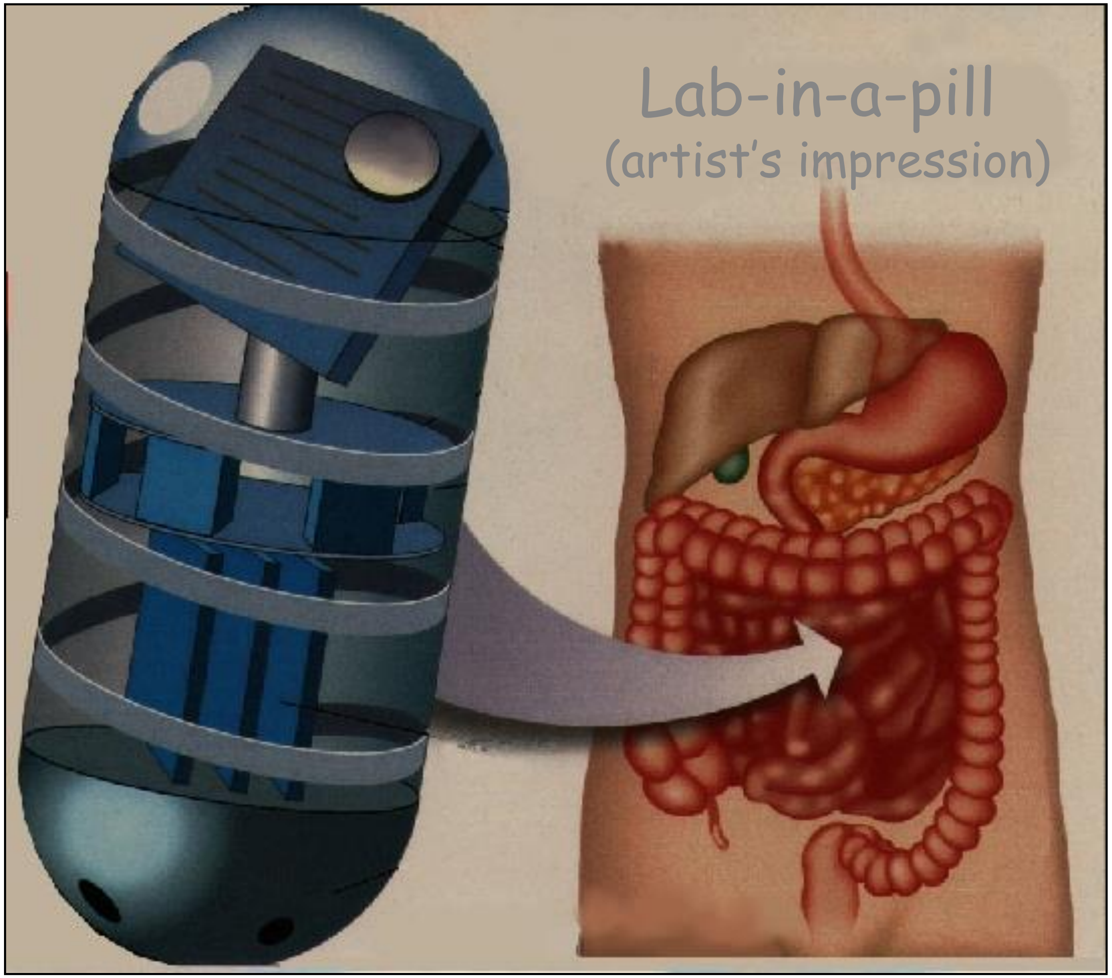
- Pashaei E, **Aydin N**, “Markovian encoding models in human splice site recognition using SVM”, *Computational Biology and Chemistry*, Vol. 73, 159-170, 2018
  - https://doi.org/10.1016/j.compbiolchem.2018.02.005
- Pashaei E, Pashaei E, Ahmady M, Ozen M, **Aydin N**, “Meta-analysis of miRNA expression profiles for prostate cancer recurrence following radical prostatectomy”, *PloS one*, Vol. 12(6), 1-23, 2017
  - https://doi.org/10.1371/journal.pone.0179543
- Pashaei E, **Aydin N**, “Binary black hole algorithm for feature selection and classification on biological data”, *Applied Soft Computing*, Vol. 56, 94-106, 2017
  - https://doi.org/10.1016/j.asoc.2017.03.002
- Pashaei E, Guzel E, Ozgurses ME, Demirel G, **Aydin N**, Ozen M, “A Meta-Analysis: Identification of Common Mir-145 Target Genes that have Similar Behavior in Different GEO Datasets”, *PloS one*, Vol. 11(9), 1-14, 2016.
  - https://doi.org/10.1371/journal.pone.0161491
- Kurt Z, **Aydin N**, Altay G, “A comprehensive comparison of association estimators for gene network inference algorithms”, *Bioinformatics*, Vol. 30(15), 2142-2149, 2014
  - https://doi.org/10.1093/bioinformatics/btu182

Examples - 2

- Mobile/miniaturized biomedical systems
  - imposes some constraint on:
    - Power consumption
    - Data rate
    - Size/area
    - Sensor technology
  - requires to employ
    - Low power architectures
    - Power aware algorithms
    - New power sources
    - High level of integration (System on Chip)

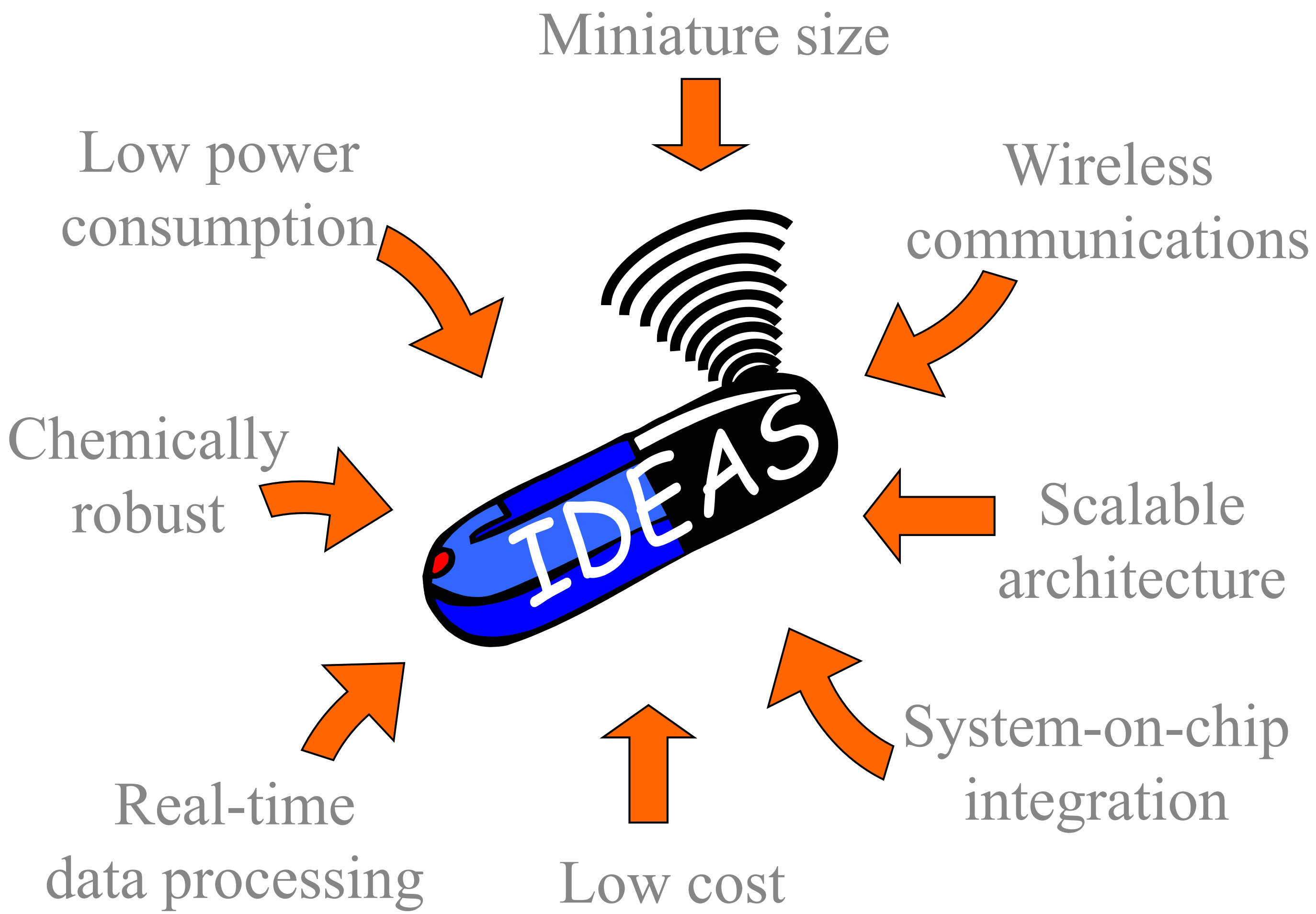
Example: Integrated Diagnostics Environmental and Analytical Systems

The concept

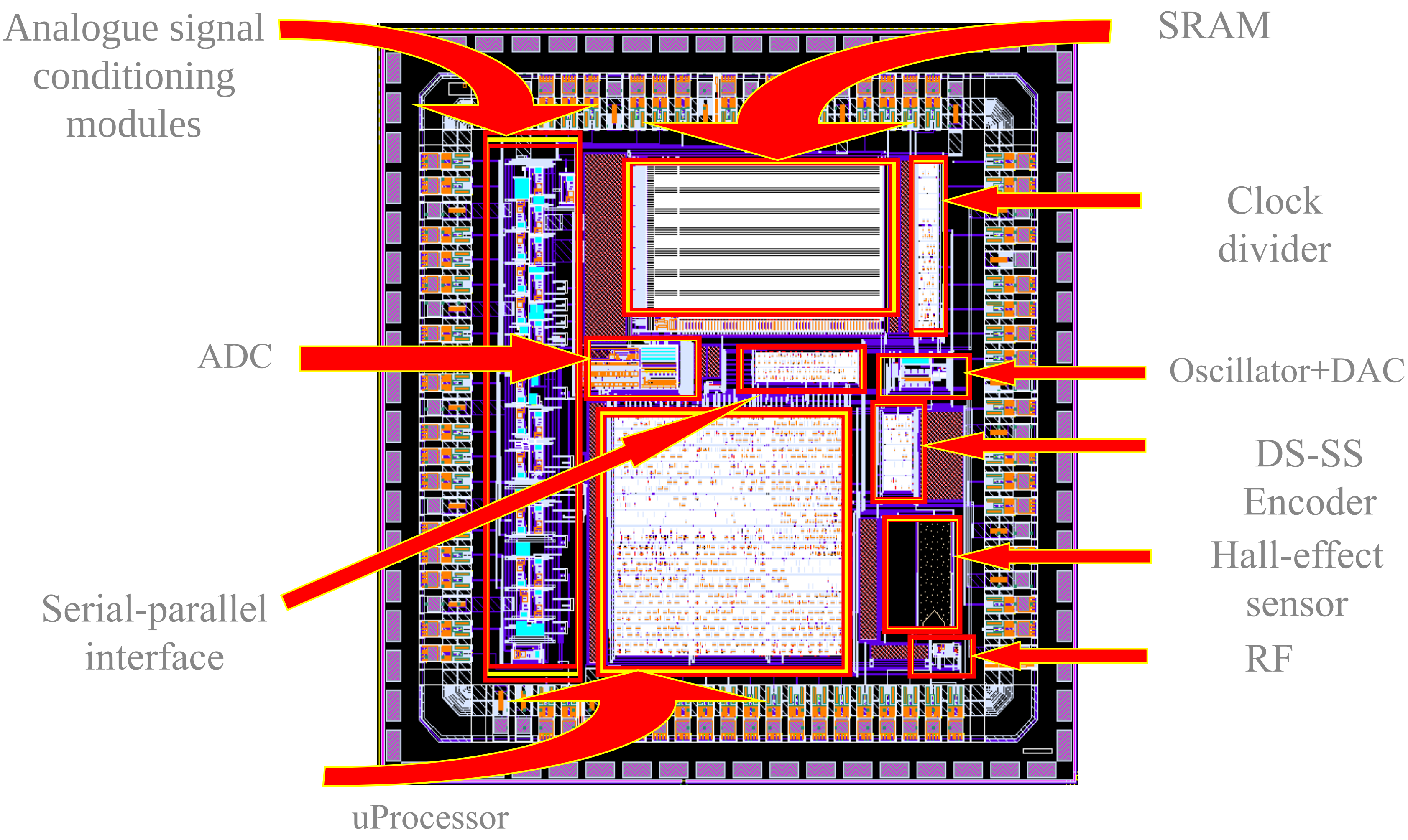


Ingestible electronic capsule integrates several sensors, amplifiers, ADC, and includes microsystem scheduling, coding and transmitting circuitry. The system is able to monitor some common physiological parameters of Gastro-intestinal tract such as temperature, pH, conductivity, and oxygen concentration

Design specifications of IDEAS



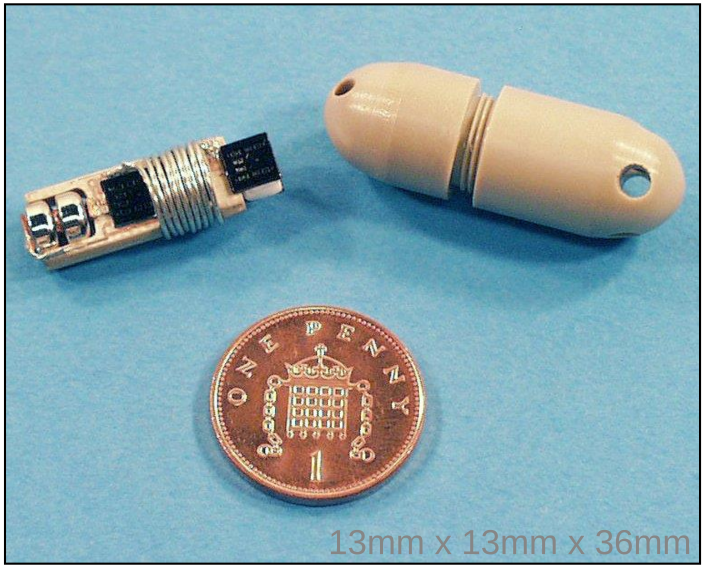
IDEAS SoC



Analogue signal conditioning modules
SRAM
Clock divider
ADC
Oscillator+DAC
DS-SS Encoder
Hall-effect sensor
Serial-parallel interface
RF
uProcessor

Protoypes

The 1st prototype



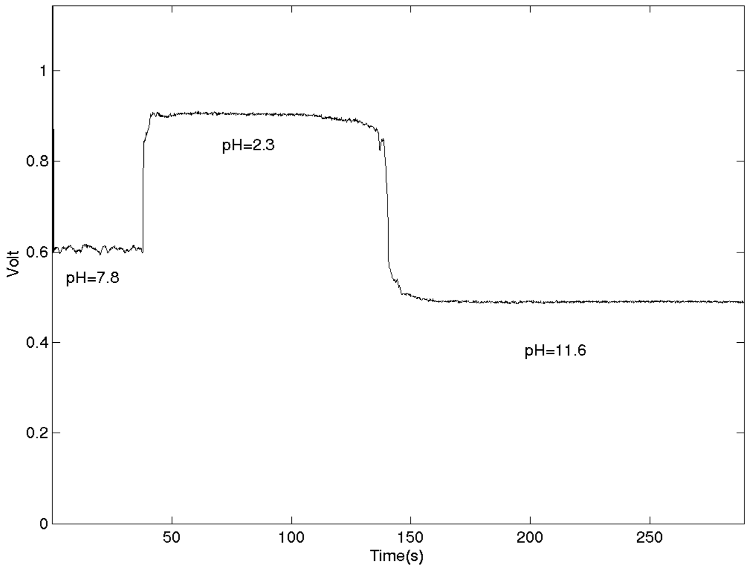
Dimensions:
32 mm (L) 11.5 mm (D)

Weight:
3.8 g (without PDMS filling)
6.4 g (with PDMS filling)

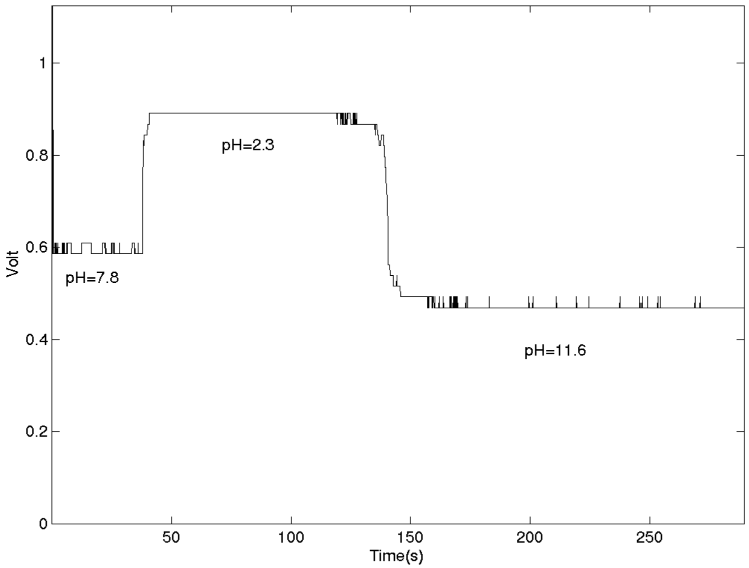
Some Results (Simulation and Real)

•1000 test data recorded from a pH sensor, which was a component of the integrated sensor microsystem, was used.
•The data were digitized by an 8 bit A/D converter modeled in Matlab and converted into a bit stream, then coded and transmitted.
Ps: The pH scale is a notation that extends from 0 to 14 with 7 as its middle point. It is a scale that measures hydrogen ion concentration

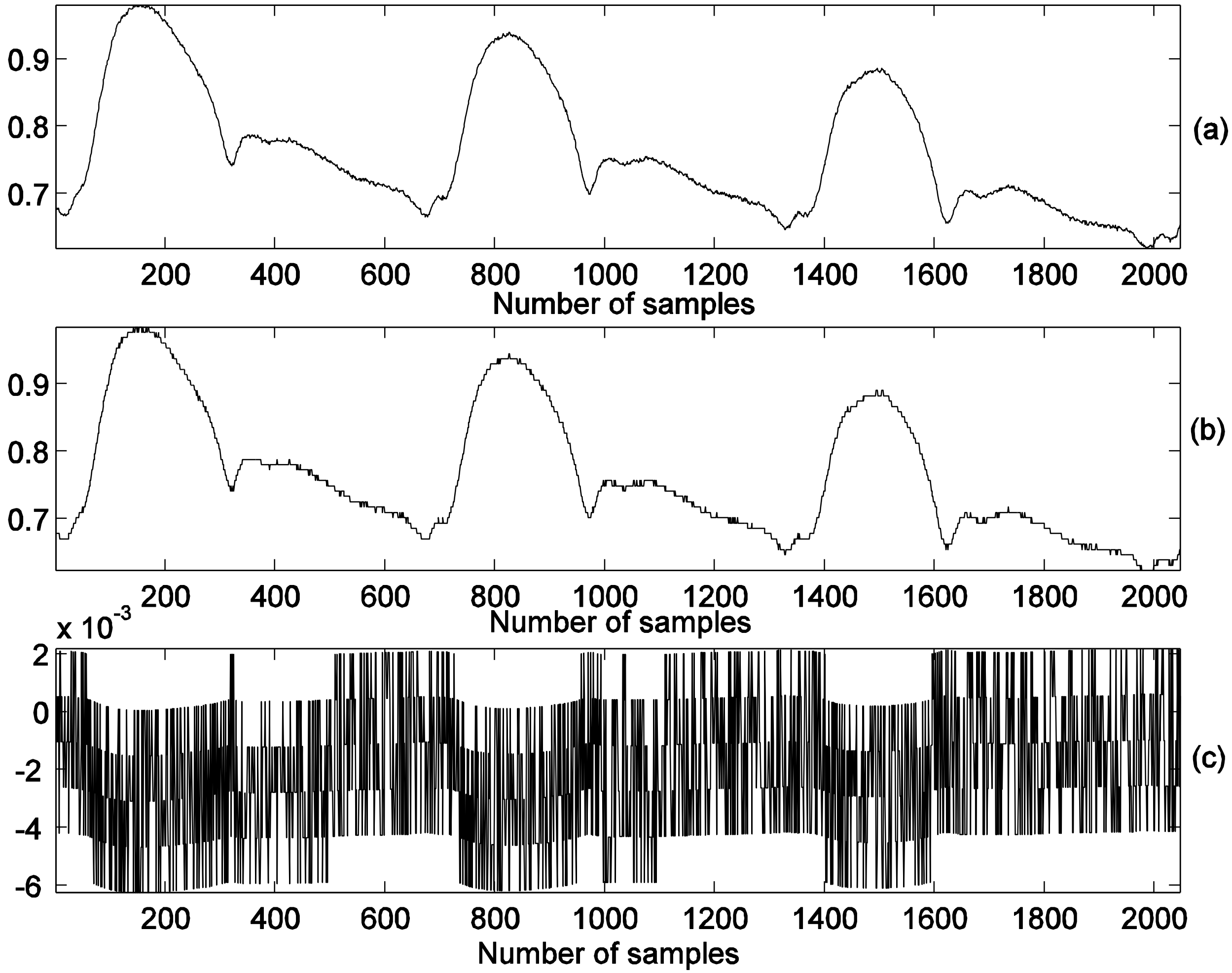
Original test data from a pH sensor



Reconstructed test data at the receiver



Some Results (Simulation and Real)



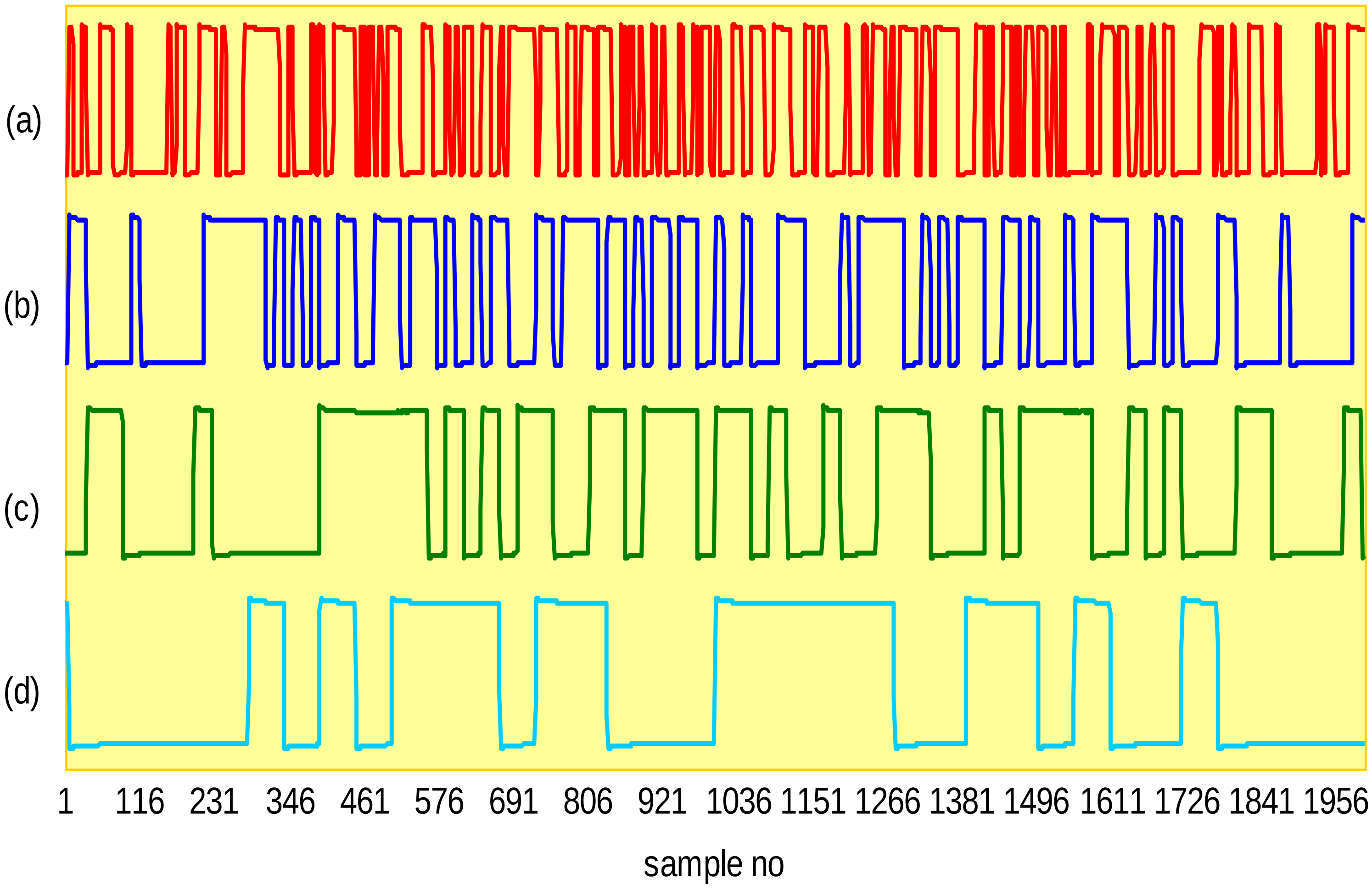
(a) Test data from an intravascular pressure sensor was encoded and transmitted by using the DS-SS TX (PN length 128). The transmitted signal was re-digitized in MATLAB using 8 bit resolution. Then it was decoded using the DS-SS receiver.

(b) Shows the reconstructed or received data at the receiver,

(c) The difference between original signal and reconstructed signal. It is possible to see the effect of digitization and integer operations

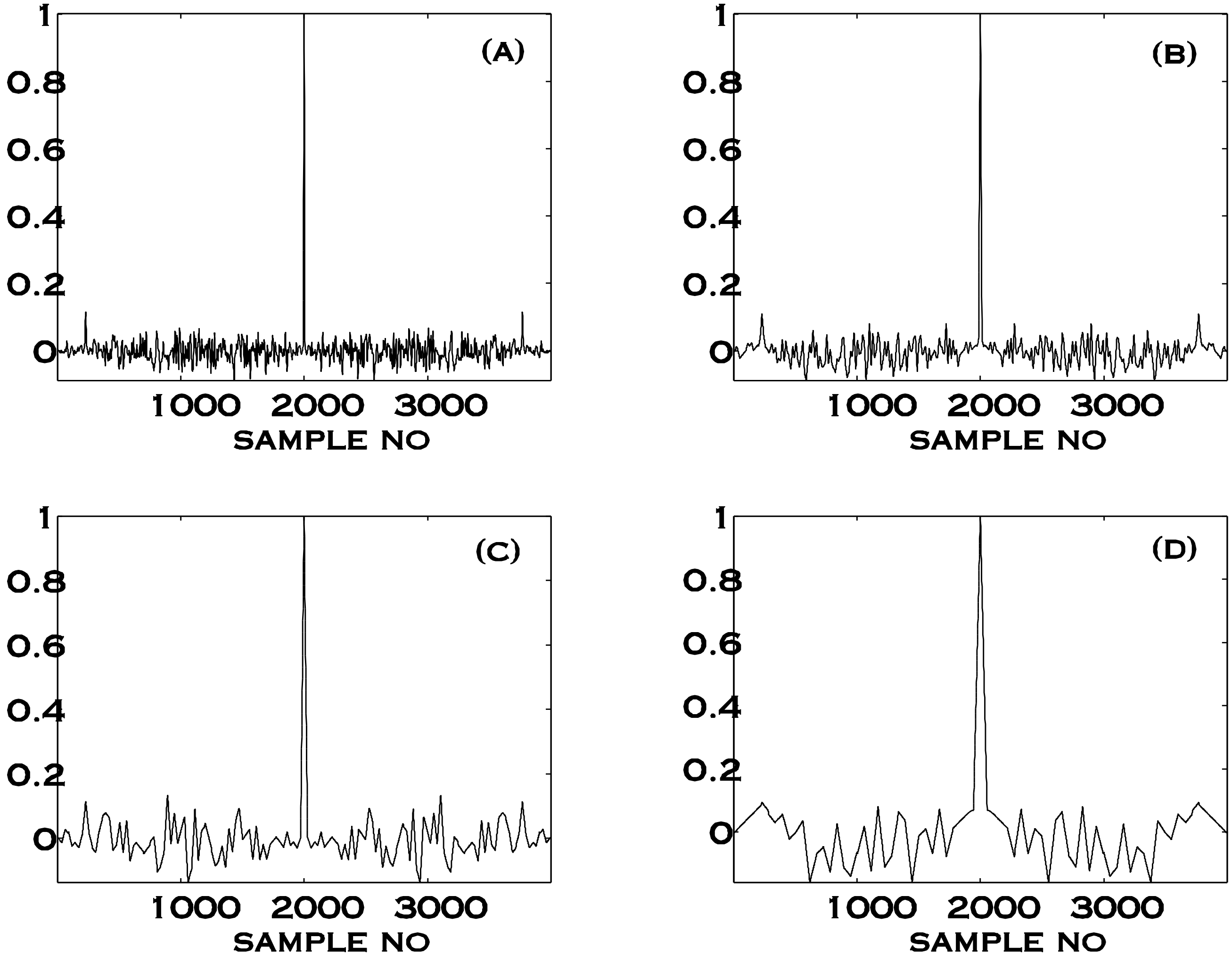
Real time test results of TX unit in IDEAS2 SoC

Test outputs of the TX recorded on an oscilloscope for the PN length of a)256, b)128, c)64, and d)32



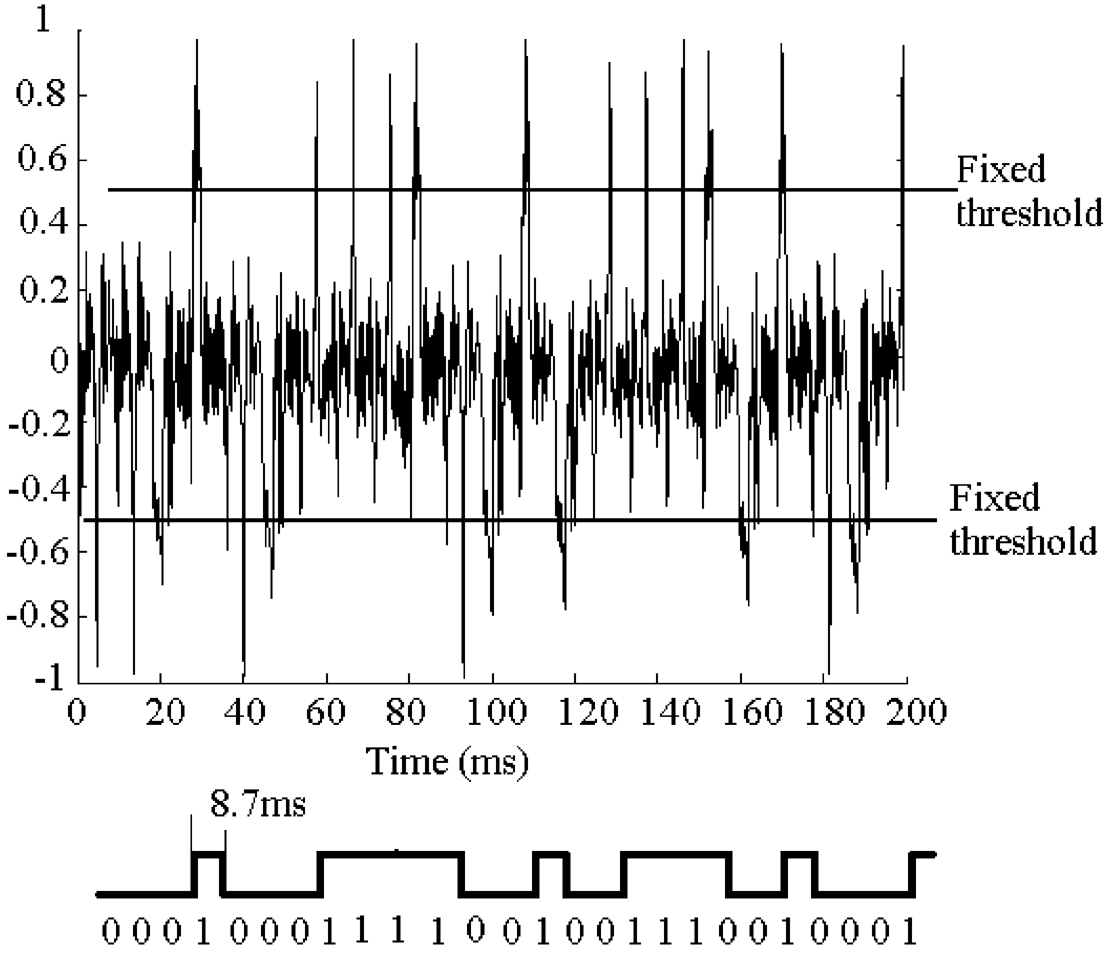
Real time test results of TX unit in IDEAS2 SoC

Corresponding normalized autocorrelation results for the PN length of a)256, b)128, c)64, and d)32



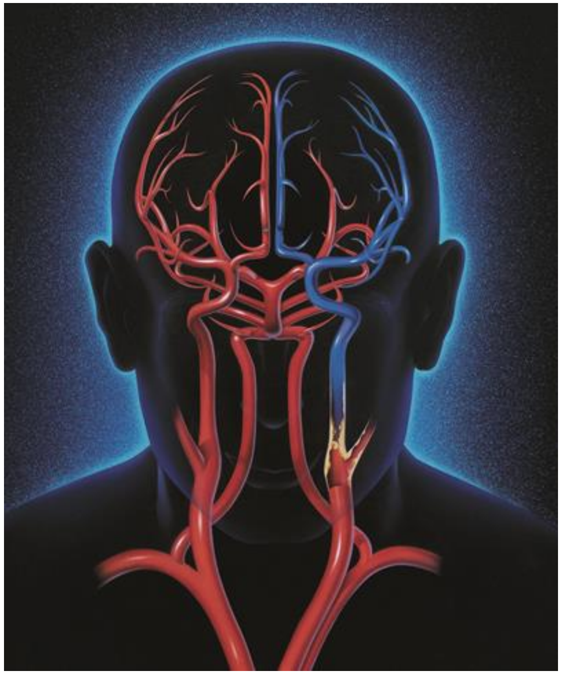
Experimental result

An experimentally obtained trace of ‘0’s and ‘1’s with value (after normalized correlation) – 1 and 1 respectively. Please note that this output was recorded while IDEAS2 chip operating, not only DS-SS encoder. This slide demonstrates that the chip in general, and DS-SS transmitter in particular functions as intended.



Examples - 3

- Asymptomatic Emboli Detection
  - Stroke is an illness causing partial or total paralysis, or death.
  - Sudden brain damage



  - Lack of blood flow to the brain caused by a clot or rupture of a blood vessel
    - The most common type of stroke (80% of all strokes) occurs when a blood vessel in or around the brain becomes plugged.

Emboli

- The "travelling clots" are called emboli.
- Solid emboli typically consist of
  - thrombus,
  - hard calcified plaque or
  - soft fatty atheroma.
- Gaseous emboli may also enter the circulation during surgery or form internally from gases that are normally dissolved in the blood.
- Any foreign body (solid or gas) that becomes free-floating in the bloodstream is called an embolus,
  - from the Greek ‘embolos’ meaning ‘a stopper’.

- Early and accurate detection of asymptomatic emboli is important in identifying patients at high risk of stroke

- They can be detected by Doppler ultrasound
  - Transcranial Doppler ultrasound (TCD)
    - 1-2 MHz

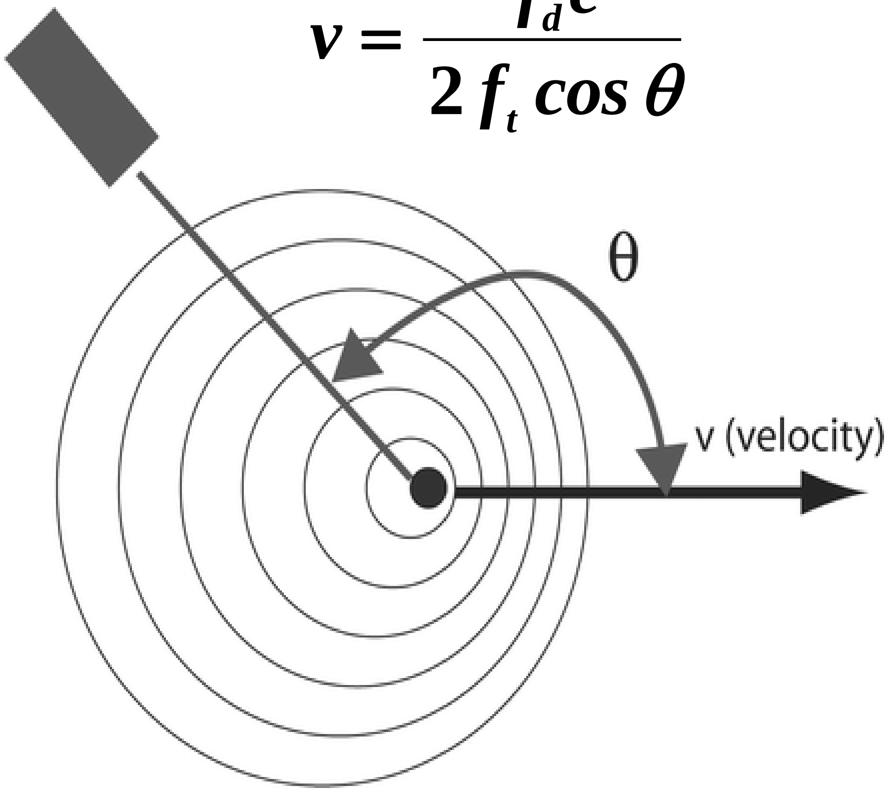
Doppler ultrasound

- A Doppler ultrasound is a noninvasive test that can be used to estimate the blood flow through your blood vessels by bouncing high-frequency sound waves (ultrasound) off circulating red blood cells.
- A regular ultrasound uses sound waves to produce images, but can't show blood flow.
- A Doppler ultrasound may help diagnose many conditions, including:
  - Blood clots
  - A blocked artery (arterial occlusion)

Doppler ultrasound

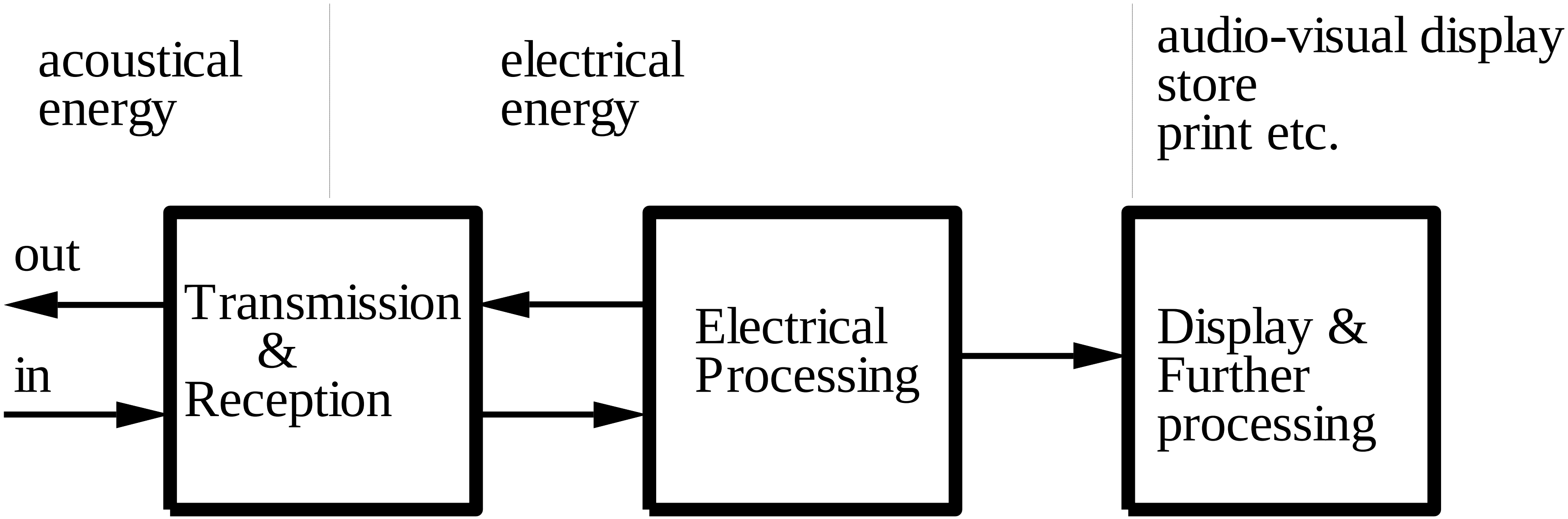
$$f_d = f_t - f_r = \frac{2vf_t \cos\theta}{c}$$

$$v = \frac{f_d c}{2 f_t \cos\theta}$$

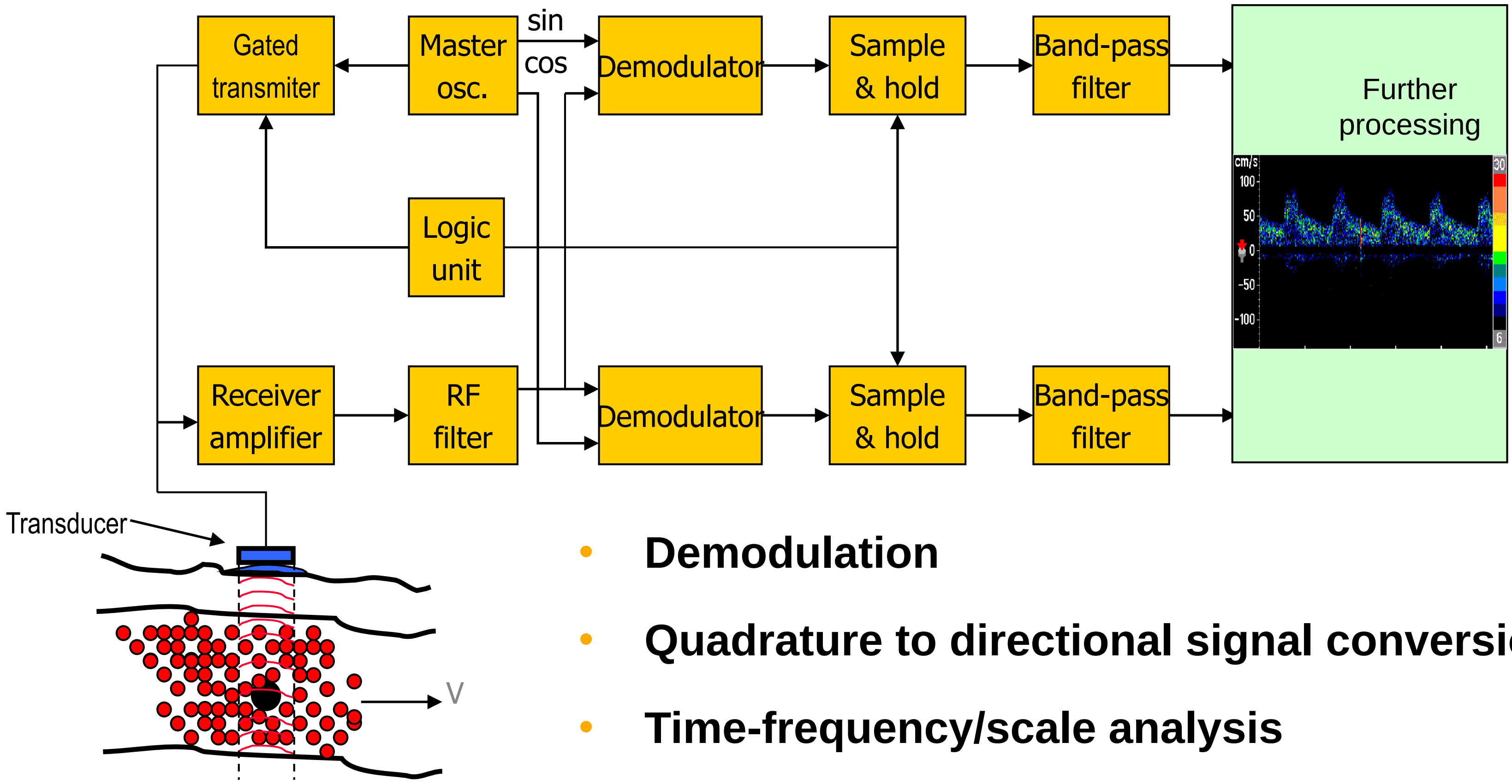


f_t is transmitted frequency
f_r is received frequency
v is the velocity of the target,
θ is the angle between the ultrasound beam and the direction of the target's motion, and
c is the velocity of sound in the medium

A general Doppler ultrasound signal measurement system

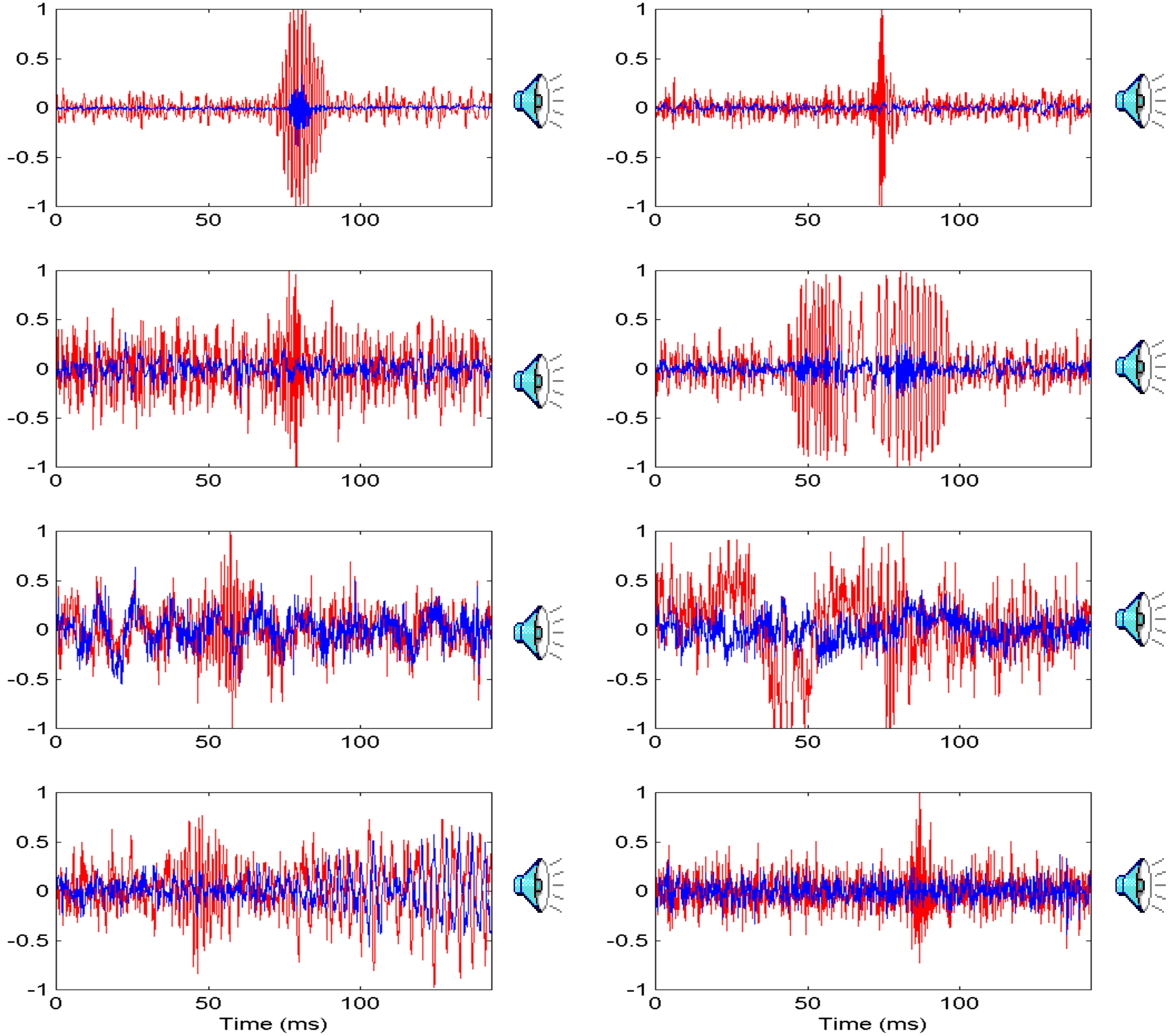


Typical Doppler System for Detecting Emboli



- **Demodulation**
- **Quadrature to directional signal conversion**
- **Time-frequency/scale analysis**
- **Data visualization**
- **Detection and estimation**
- **Derivation of diagnostic information**

Examples of Embolic Signals

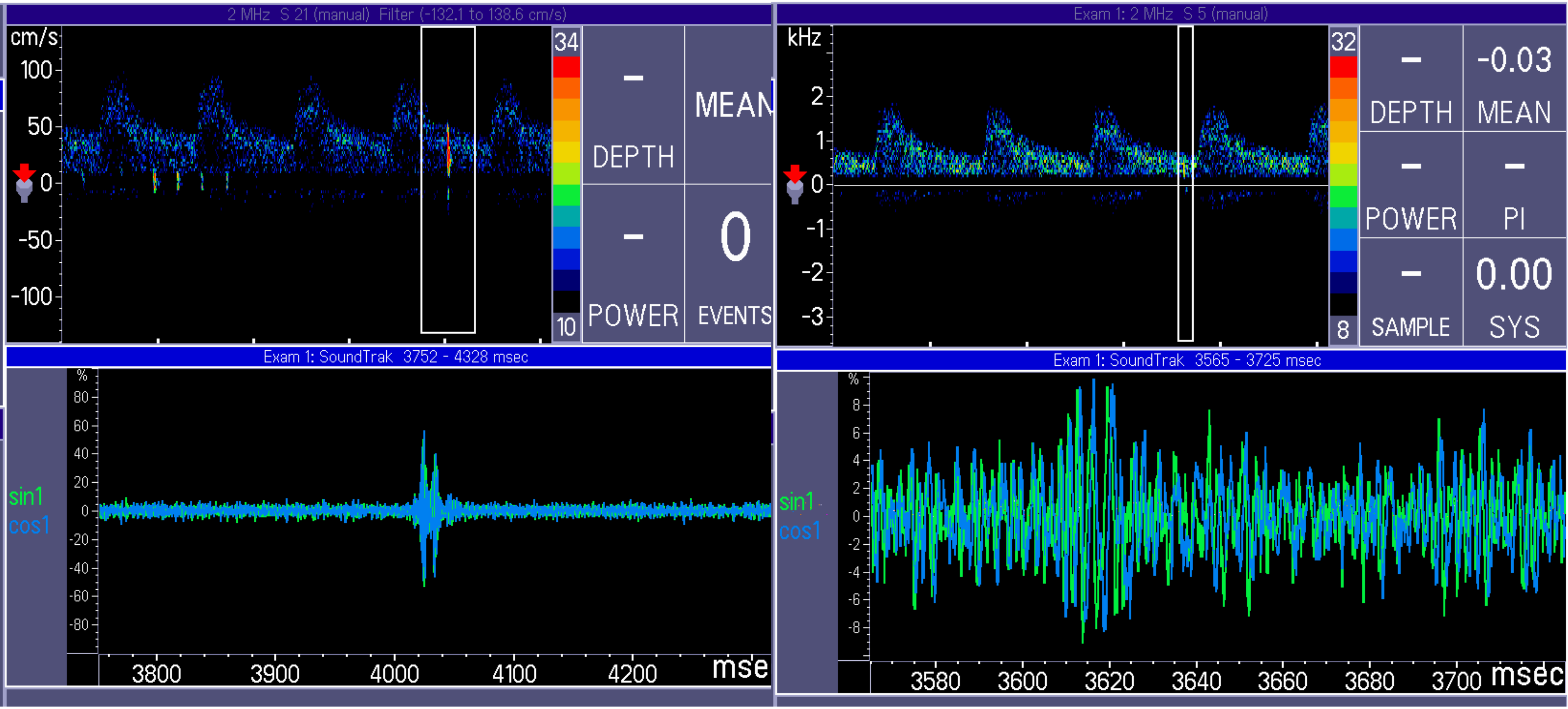


(red: forward,

blue: reverse)

More examples

Sonogram Display of Embolic Signals



TF and TS analysis

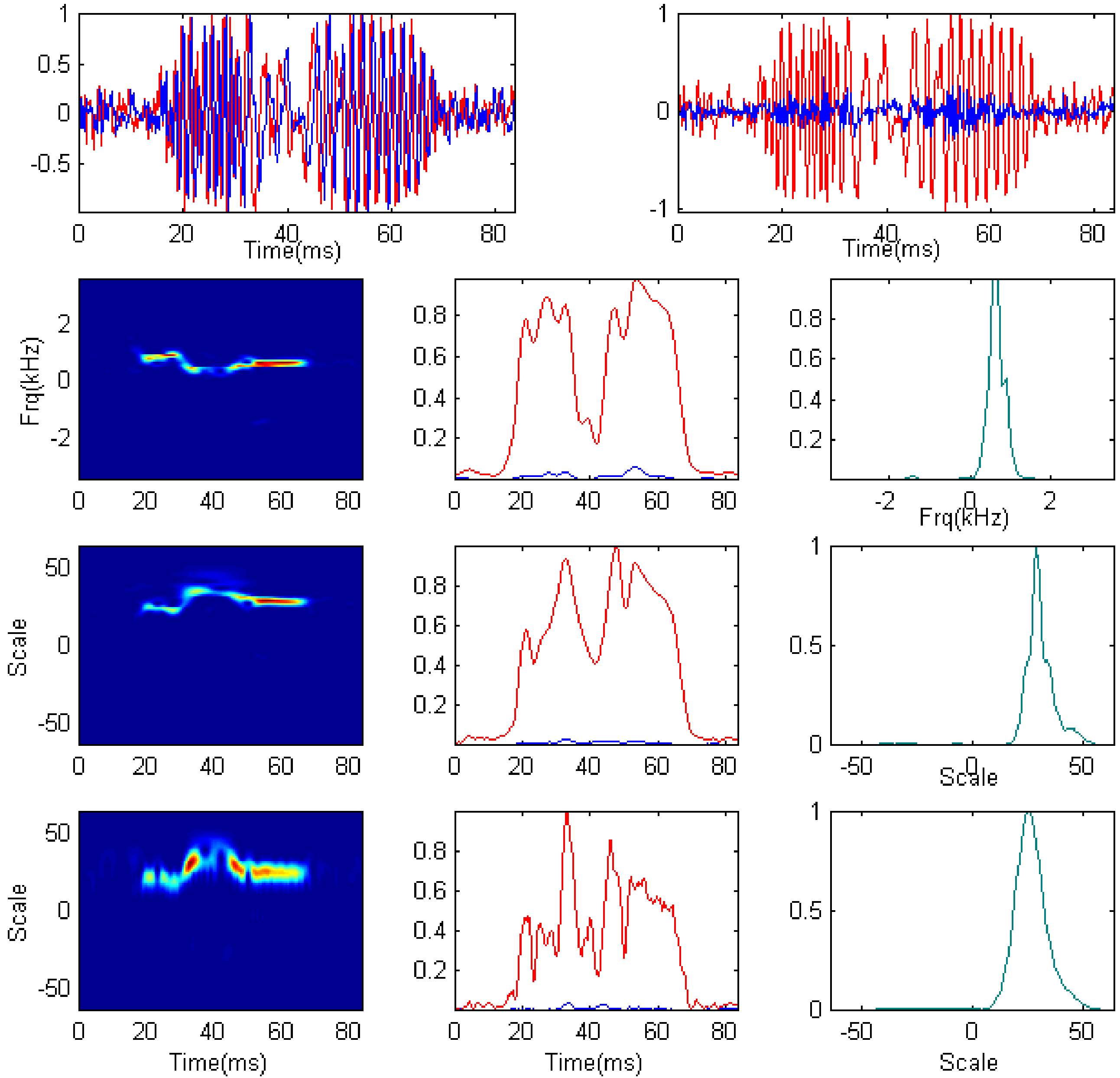
$$F_s(t,v) = \int_{-\infty}^{+\infty} s(\tau) g^*(\tau - t) e^{-j2\pi v\tau} d\tau$$

$$W_s(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} s(t) \psi^*\left(\frac{t-b}{a}\right) dt$$

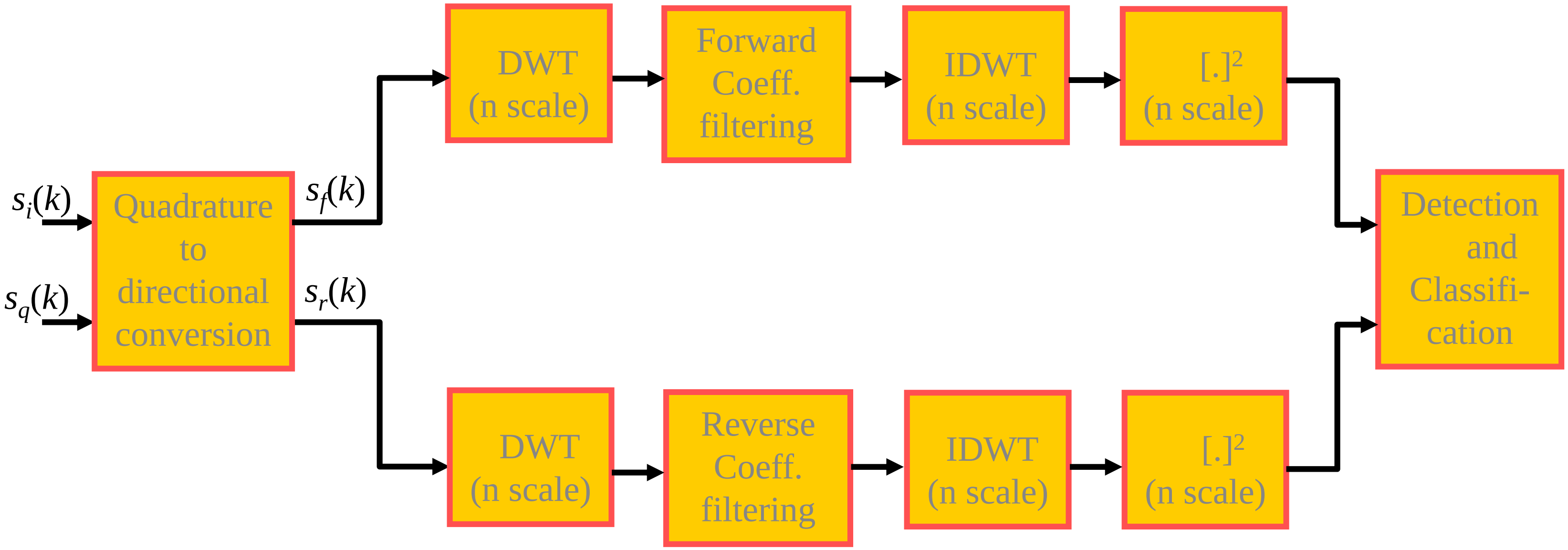
- TF tiling is linear
- Decomposes a time series into TF space
- Trade-off between frequency and time resolution
- Assumes the signal is stationary within the analysis window
- Fast algorithms exist

- TS tiling is logarithmic
- Decomposes a time series into TS space
- TF resolution compromise is optimised
- Ideal for the analysis of sudden short duration signal changes
- Fast algorithms exist

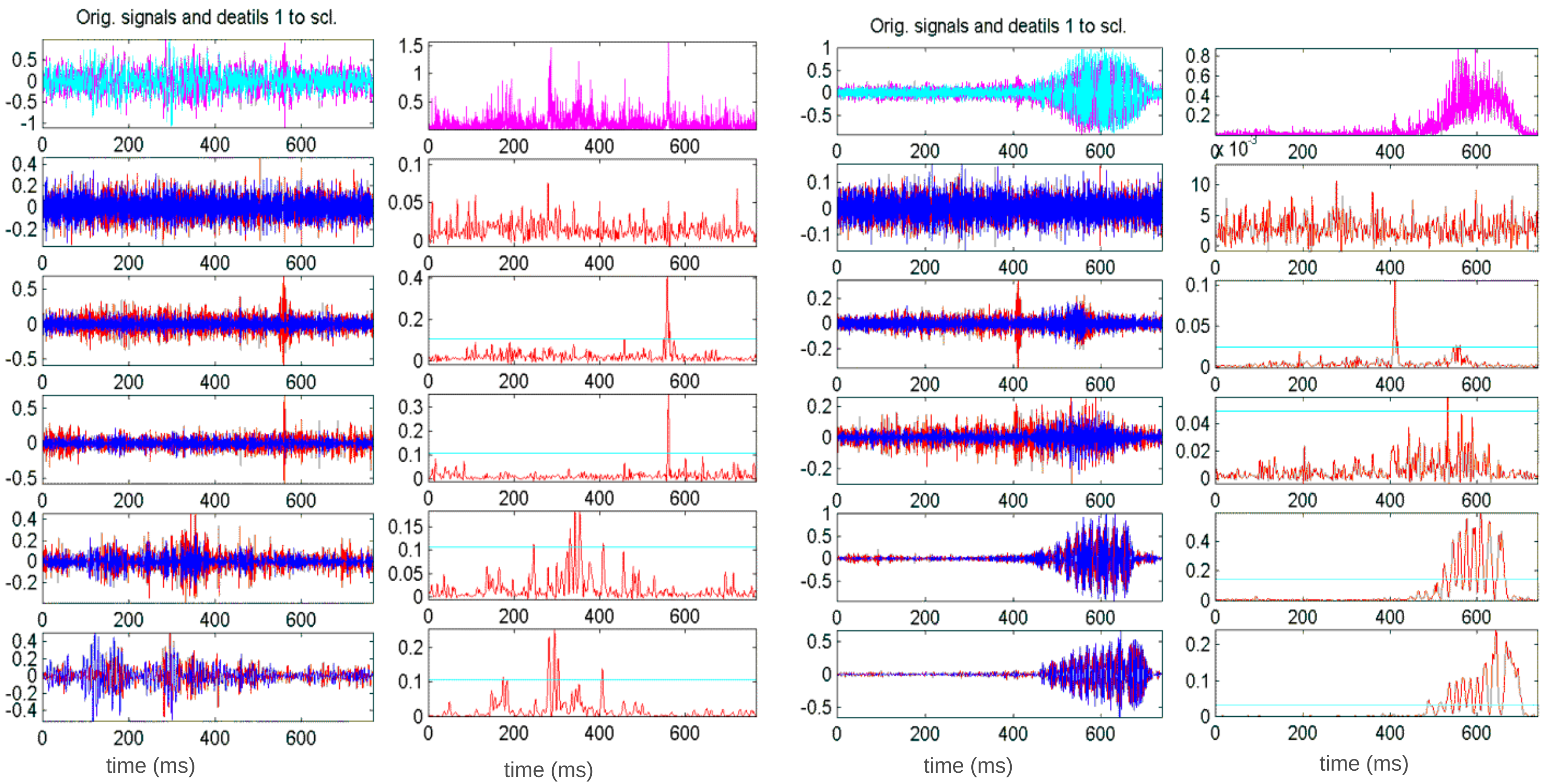
Example: analysis of Embolic Doppler signal



Detection of Embolic Signals Using DWT and Fuzzy Logic

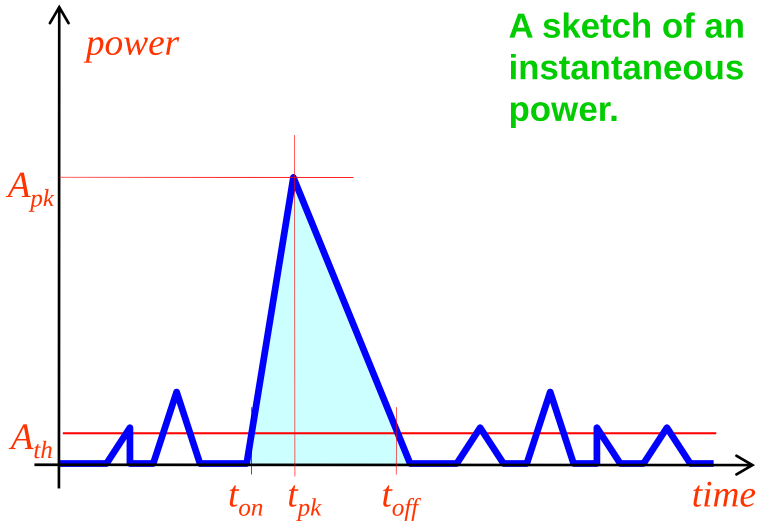


DWT of ES



Orig. signals and deatils 1 to scl.
Orig. signals and deatils 1 to scl.
time (ms)
time (ms)
time (ms)
time (ms)

Some Parameters Used in Detection



A_{th} : threshold value

P2TR: peak value to threshold ratio

TP2TR: total peak value to threshold ratio

RR : rise rate

FR : fall rate

F2RM : peak forward power to reverse power ratio

TF2R : total forward power to reverse power ratio

$$A_{th} = \sigma_{fn}\sqrt{\log_2 N} + \sigma_{rn}\sqrt{\log_2 N}$$

$$P2TR = 10\log\frac{A_{pk}}{A_{th}} \quad (dB)$$

$$TP2TR = 10\log\frac{\sum_{k=t_{on}}^{t_{off}} A_f(k)}{A_{th}} \quad (dB)$$

$$RR = \frac{P2TR}{t_{pk} - t_{on}} \quad (dB/ms)$$

$$FR = \frac{P2TR}{t_{off} - t_{pk}} \quad (dB/ms)$$

$$F2RM = 10\log\frac{A_f(t_{pk})}{A_r(t_{pk})} \quad (dB)$$

$$TF2R = 10\log\frac{\sum_{k=t_{on}}^{t_{off}} A_f(k)}{\sum_{k=t_{on}}^{t_{off}} A_r(k)} \quad (dB)$$

Some Parameters Used in Detection

These parameters are based on the narrow-band assumption

t_s : averaged time centre of the signal

f_s : averaged frequency centre of the signal

T_s^2: time spreading

B_s^2: frequency spreading

$a(t)$: instantaneous amplitude

$f(t)$: instantaneous frequency

$s_a(t)$: complex quadrature signal given as

$$s_a(t)=s(t)+jH\{s(t)\}$$

Where $H\{s(t)\}$ is Hilbert transform of $s(t)$

$$t_s = \frac{1}{E_s}\int_{-\infty}^{+\infty} t\left|s(t)\right|^2 dt$$

$$f_s = \frac{1}{E_s}\int_{-\infty}^{+\infty} f\left|S(f)\right|^2 df$$

$$T_s^2 = \frac{1}{E_s}\int_{-\infty}^{+\infty} (t-t_s)^2\left|s(t)\right|^2 dt$$

$$B_s^2 = \frac{1}{E_s}\int_{-\infty}^{+\infty} (f-f_s)^2\left|S(t)\right|^2 df$$

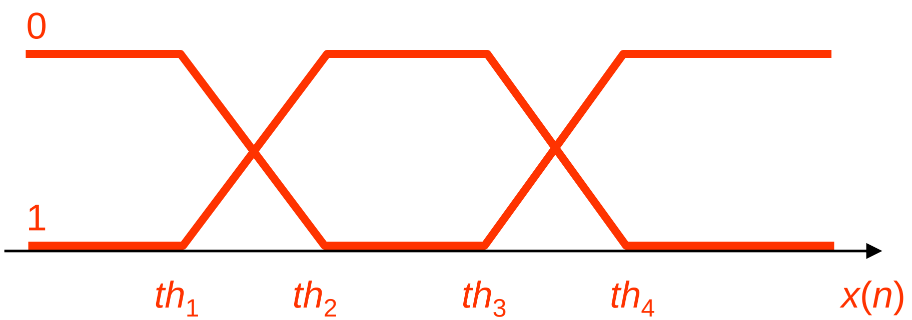
$$E_s = \int_{-\infty}^{+\infty}\left|s(t)\right|^2 dt \quad < +\infty$$

$$a(t) = \left|s_a(t)\right|$$

$$f(t) = \frac{1}{2\pi}\frac{d \arg s_a(t)}{dt}$$

Fuzzy Membership Function and Detection Rules

Trapezoidal membership function used for the derivation of membership values



Membership value (MV):

If $x(n) < th_1$; MV $\Rightarrow$ 0 or 1

If $x(n) \geq th_1$ & $\leq th_2$;

MV $\Rightarrow$ $Z_1 = \frac{x(n) - th_1}{th_2 - th_1}, \; Z_3 = 1 - Z_1$

If $x(n) > th_2$ & $< th_3$;

MV $\Rightarrow$ 0 or 1

If $x(n) \geq th_3$ & $\leq th_4$;

MV $\Rightarrow$ $Z_2 = \frac{x(n) - th3}{th_4 - th_3}, \; Z_4 = 1 - Z_2$

If $x(n) > th_4$; MV $\Rightarrow$ 0 or 1

MEMBERSHIP RULES FOR THE PARAMETERS

$x(n)$	$< th_1$			$\geq th_1$ & $\leq th_2$			$> th_2$ & $< th_3$			$\geq th_3$ & $\leq th_4$			$> th_4$		
	ar	em	sp	ar	em	sp	ar	em	sp	ar	em	sp	ar	em	sp
$P2TR$	0	0	1	0	z_1	$1-z_1$	0	1	0	z_2	$1-z_2$	0	1	0	0
$TP2TR$	0	0	1	0	z_1	$1-z_1$	0	1	0	z_2	$1-z_2$	0	1	0	0
$F2RM$	1	0	0	$1-z_1$	0	z_1	0	0	1	0	z_2	$1-z_2$	0	1	0
$TF2R$	1	0	0	$1-z_1$	0	z_1	0	0	1	0	z_2	$1-z_2$	0	1	0
RR	1	0	0	$1-z_1$	z_1	0	0	1	0	0	$1-z_2$	z_2	0	0	1
FR	1	0	0	$1-z_1$	z_1	0	0	1	0	0	$1-z_2$	z_2	0	0	1
t_s	0	0	1	0	$1-z_1$	z_1	0	1	0	z_2	$1-z_2$	0	1	0	0
f_s	0	0	1	0	z_1	$1-z_1$	0	1	0	z_2	$1-z_2$	0	1	0	0
T_s^2	0	0	1	0	$1-z_1$	z_1	0	1	0	z_2	$1-z_2$	0	1	0	0
B_s^2	1	0	0	$1-z_1$	z_1	0	0	1	0	0	$1-z_2$	z_2	0	0	1
VIE	0	0	1	z_1	0	$1-z_1$	1	0	0	$1-z_2$	z_2	0	0	1	0
VIF	1	0	0	$1-z_1$	z_1	0	0	1	0	0	$1-z_2$	z_2	0	0	1

ar: artifact, em: emboli, sp: speckle; $z_1 = (th_2 - th_1)^{-1}(x(n) - th_1)$, $z_2 = (th_4 - th_3)^{-1}(x(n) - th_3)$

THRESHOLD VALUES OF THE PARAMETERS USED IN THE DETECTION

	th_1	th_2	th_3	th_4
$P2TR$ (dB)	6	12	14	20
$TP2TR$ (dB)	17	23	26	38
$F2RM$ (dB)	10	20	22	26
$TF2R$ (dB)	4	8	10	20
RR (ms)	0.6	1.4	2	5
FR (ms)	0.6	1.4	2	6
t_s (ms)	10	20	60	120
f_s/F_s (unit)	0.01	0.035	0.08	0.1
T_s^2 (ms^2)	6	18	40	100
B_s^2/F_s (unit)	0.03	0.06	0.1	0.4
VIE (unit)	12	60	100	140
VIF/F_s (unit)	0.008	0.016	0.021	0.04

F_s= Sampling frequency

MEAN AND STANDARD DEVIATIONS OF SOME PARAMETERS FOR EMBOLIC SIGNALS, ARTIFACTS, AND DS

	ES		Artifact		DS	
	Mean	SD	Mean	SD	Mean	SD
SMP	2.45	0.83	5.96	0.77	3.27	0.72
$TP2TR$	25.13	3.17	34.57	6.62	18.32	3.13
$P2TR$	12.1	3.01	14.16	6.67	6.29	1.82
$F2RM$	25.44	7.01	9.4	11.15	23.18	7.59
$F2R8$	1.29	7.13	1.79	11.25	-0.58	6.18
RR	3.88	2.53	0.95	0.59	3.52	2.05
FR	4.65	3.15	0.97	0.63	4.00	2.12
$TF2R$	15.30	6.24	4.05	9.82	14.13	4.03
t_s	51.59	31.7	146.38	76.8	6.96	4.9
f_s	0.114	0.049	0.014	0.007	0.108	0.189
T_s^2	73.87	43.3	185.08	106.6	10.86	8.6
B_s^2	0.08	0.038	0.033	0.031	0.492	0.438
VIE	101.34	96.1	43.47	63.3	12.3	5.6
VIF	0.016	0.008	0.012	0.021	0.021	0.012

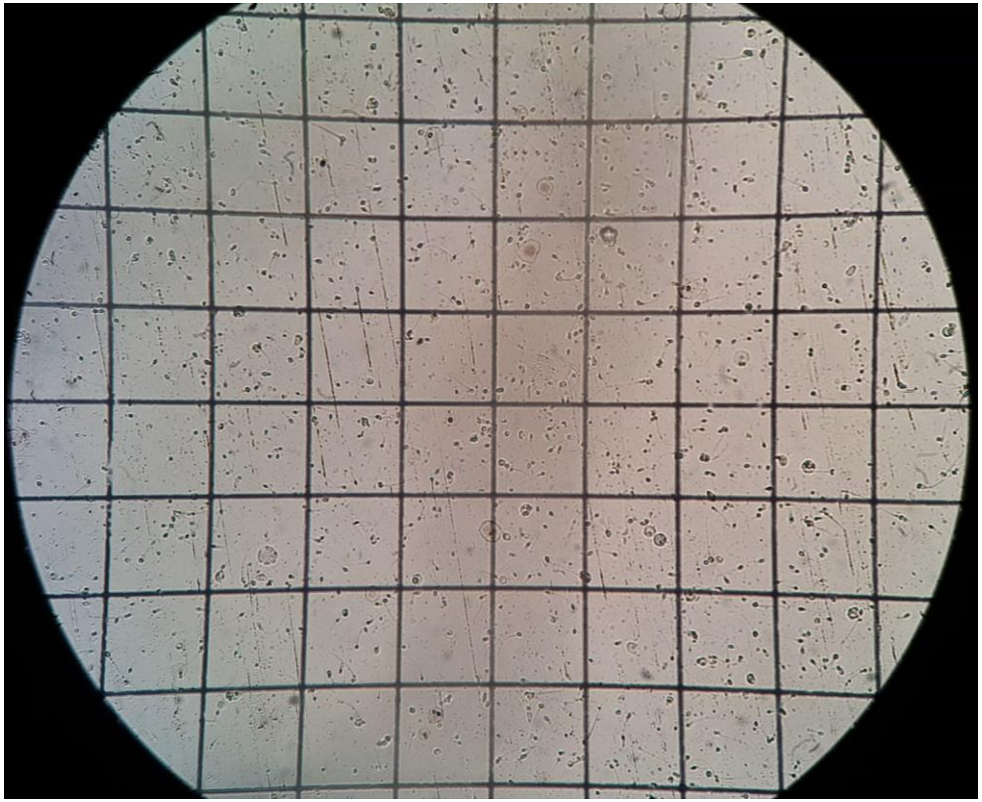
Detection Results

Data set 1	Data set 2
100 ES 98% as embolic signal 1% as artifact 1% disputed	100 ES 95% as ES 3% as artifact 2% disputed
100 artifacts 96% detected as artifact 4% disputed	100 artifacts 98% as artifact 2% as ES
100 DS 93% as DS 6% as ES 1% disputed	100 DS 95% as DS 1% as artifact 4% disputed

When it was tested on a third data set, 198 ES out of 202 were detected as ES

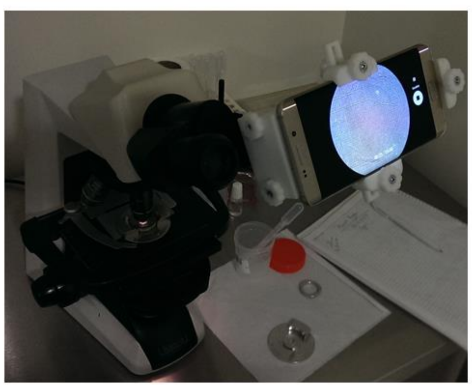
Examples - 4

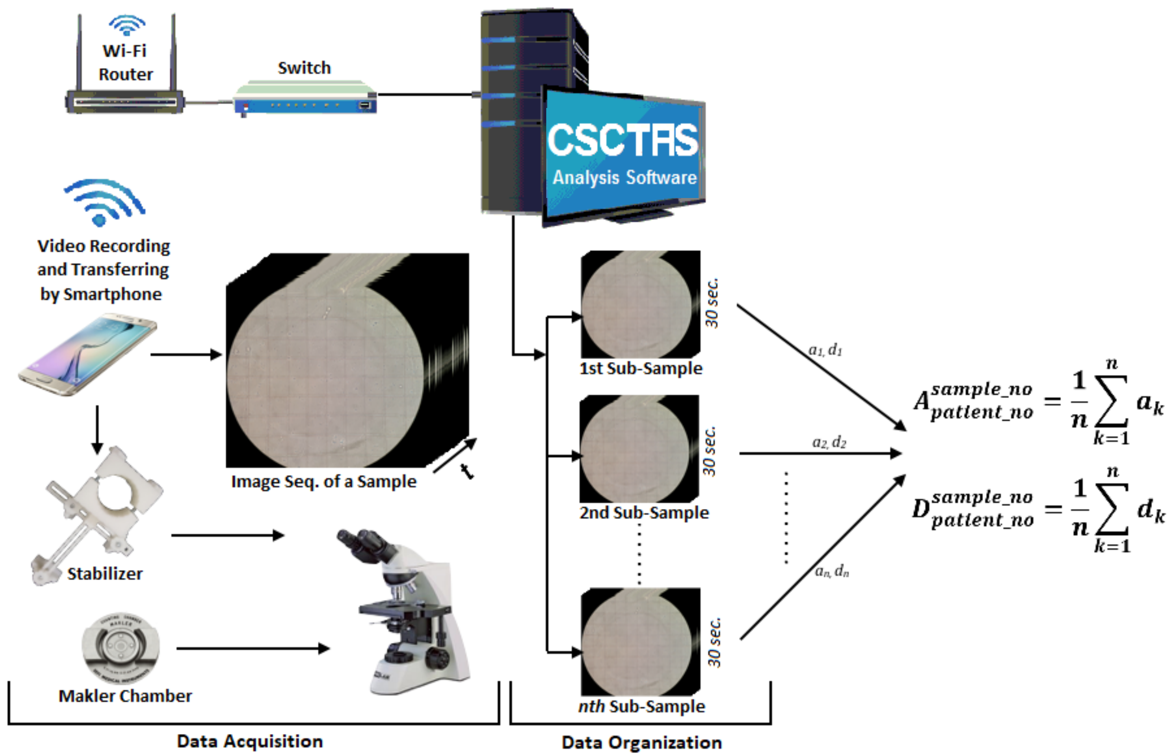
- Smartphone Based Computerized Sperm Analysis



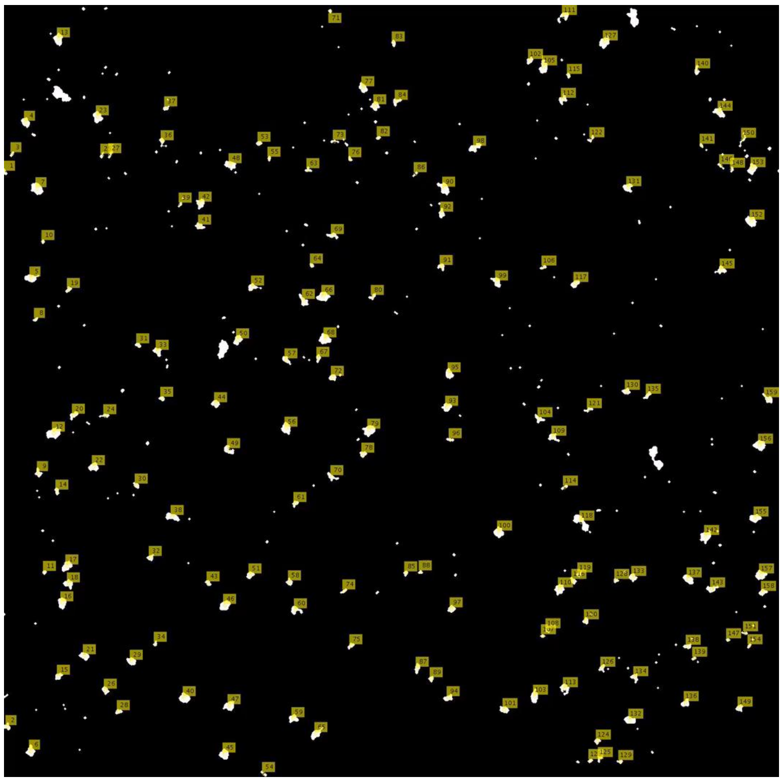
  - Sperm Concentration Analysis
  - Sperm Morphology Analysis
  - Sperm Motility Analysis

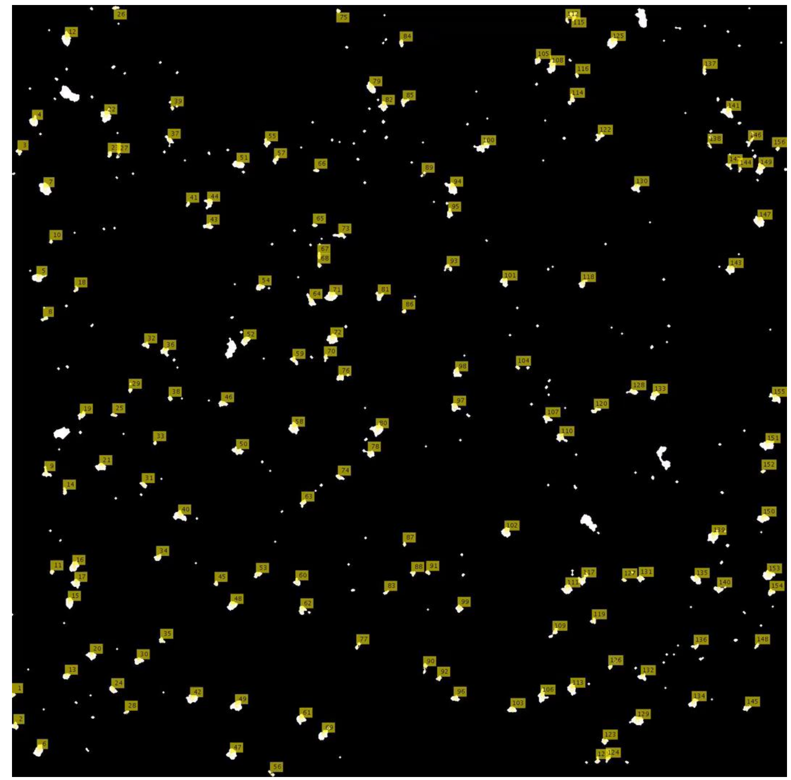
Smartphone Based Data Acquisition Approach





Feature Matching Based Video Stabilization



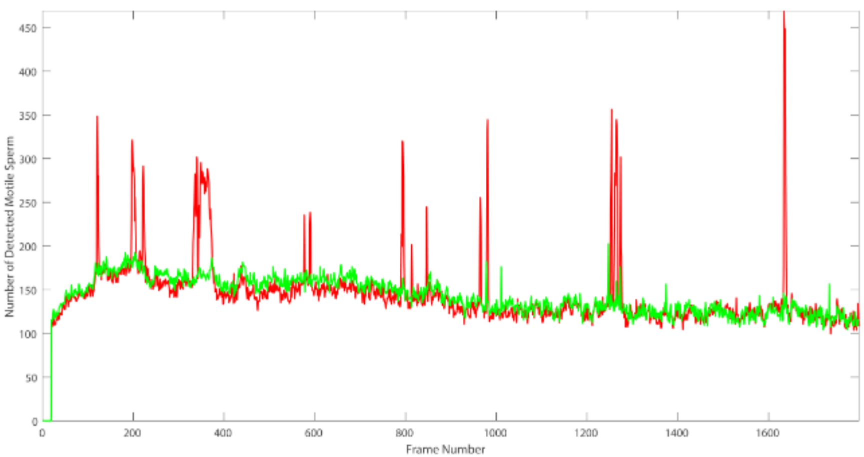


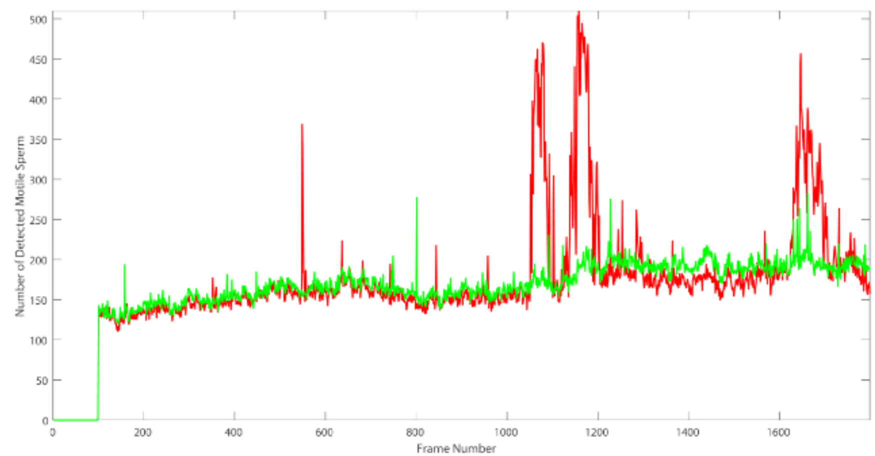
*Hamza Osman Ilhan, Nizamettin Aydin, A novel data acquisition and analyzing approach to spermiogram tests, Biomedical Signal Processing and Control, Volume 41, March 2018, Pages 129-139

Feature Matching Based Video Stabilization

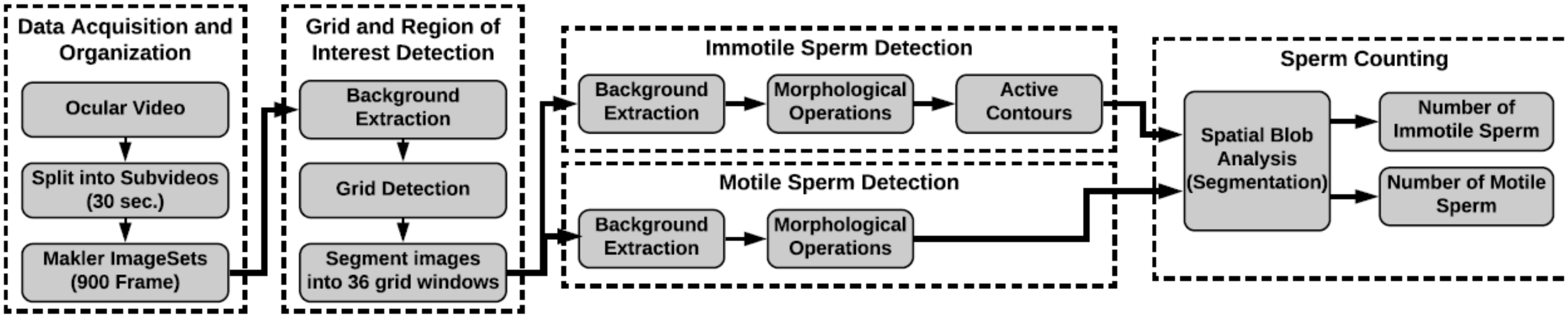
Table 3.2.1 Total Number of Vibrated Frame Numbers in Sequence

Frame Countings		Normospermia								Oligospermia				Azoospermia	
		Subj1	Subj2	Subj3	Subj4	Subj5	Subj6	Subj7	Subj8	Subj1	Subj2	Subj3	Subj4	Subj1	Subj2
Original	Vid1	53	24	6	28	280	35	132	188	282	88	34	8	48	28
	Vid2	105	50	8	150	47	457	52	114	157	163	60	57	257	57
	Vid3	39	9	26	66	364	188	54	87	105	238	84	16	173	45
	Minimum	39	9	6	28	47	35	52	87	105	88	34	8	48	28
SURF	Vid1	12	5	1	12	92	4	3	18	12	21	3	18	25	0
	Vid2	150	43	0	9	0	128	2	18	8	4	7	5	88	0
	Vid3	10	7	3	15	16	39	38	12	13	55	11	0	34	0
	Minimum	10	5	0	9	0	4	2	12	8	4	3	0	25	0





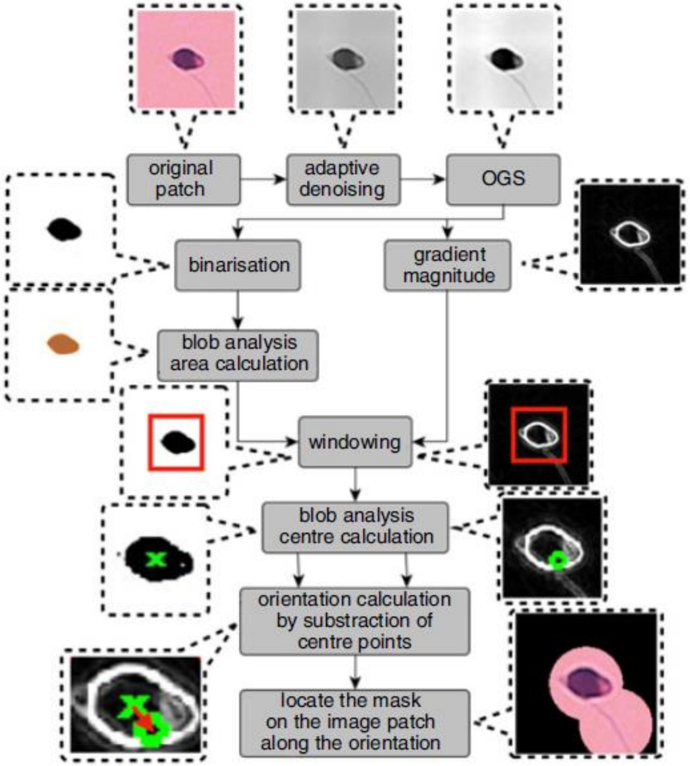
Sperm Concentration Analysis



$million/ml$		Motile Sperm Concentration						Immotile Sperm Concentration					
		VA		CASA		CSCS		VA		CASA		CSCS	
		Avg.	$\pm\sigma$	Avg.	$\pm\sigma$	Avg.	$\pm\sigma$	Avg.	$\pm\sigma$	Avg.	$\pm\sigma$	Avg.	$\pm\sigma$
Subjects	#1	0	0	No Operation		0	0	0	0	No Operation		0.28	0.13
	#2	0.25	0.43			0.3	0	3.25	0.83			3	0.41
	#3	5.25	0.83	5.1	0.17	4.67	0.56	8.25	0.43	9	0.16	7.53	0.2
	#4	10.25	0.43	10	0.21	9	0.64	11.25	0.43	12.2	0.59	9.93	0.63
	#5	2.25	0.43	1.7	0.24	2.33	0.65	23.75	2.59	26.6	0.90	23.65	2.65
	#6	21.75	0.83	22.2	0.66	21.38	0.94	11.75	1.3	12.1	0.32	10.68	0.58
	#7	20.5	1.12	21.1	0.85	19.6	0.97	23.5	1.12	23	0.57	22.65	0.92
	#8	24	0.71	24.4	0.36	23.15	0.56	25.5	0.87	26.4	0.51	24.4	0.58

*Hamza Osman Ilhan, Nizamettin Aydin, Smartphone based sperm counting - an alternative way to the visual assessment technique in sperm concentration analysis, Multimedia Tools And Applications, (Accepted - In Publication)

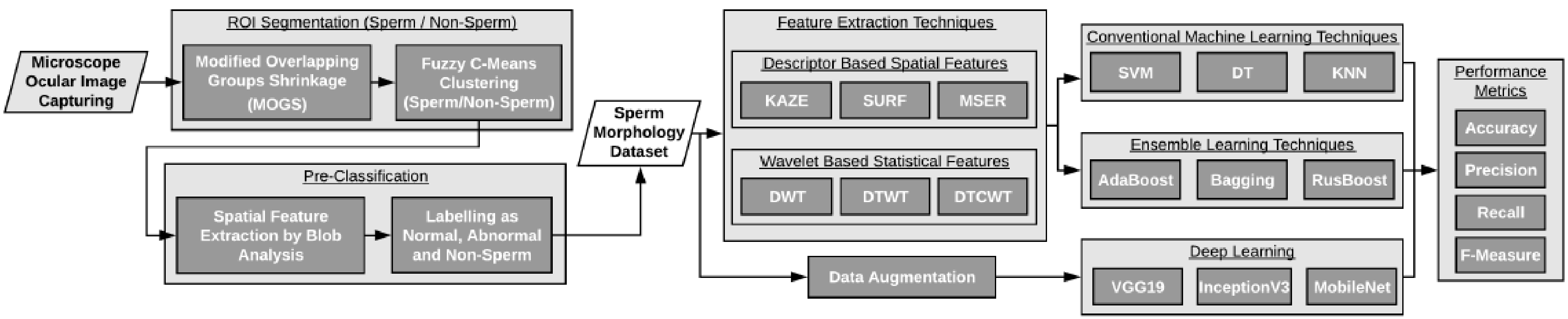
Sperm Morphology Analysis



Sperm shapes	No. of sperms	No. of correct masking, %	k-NN image classification ($k = 5$) (precision–sensitivity–accuracy), %		
			Original (raw)	k-means segmented	Masked by our method
Normal	59	57 (96.6)	96–63–73	71–64–81	79–63–82
Tapered	55	51 (92.7)	63–45–67	51–46–72	52–52–77
Pyriform	58	55 (94.8)	8–25–72	17–66–76	54–63–80
Amorphous	59	51 (86.4)	11–48–76	78–54–78	43–54–76
Total	231	216 (93.5)	—		
Overall accuracy, %			44.4	53.7	57.4

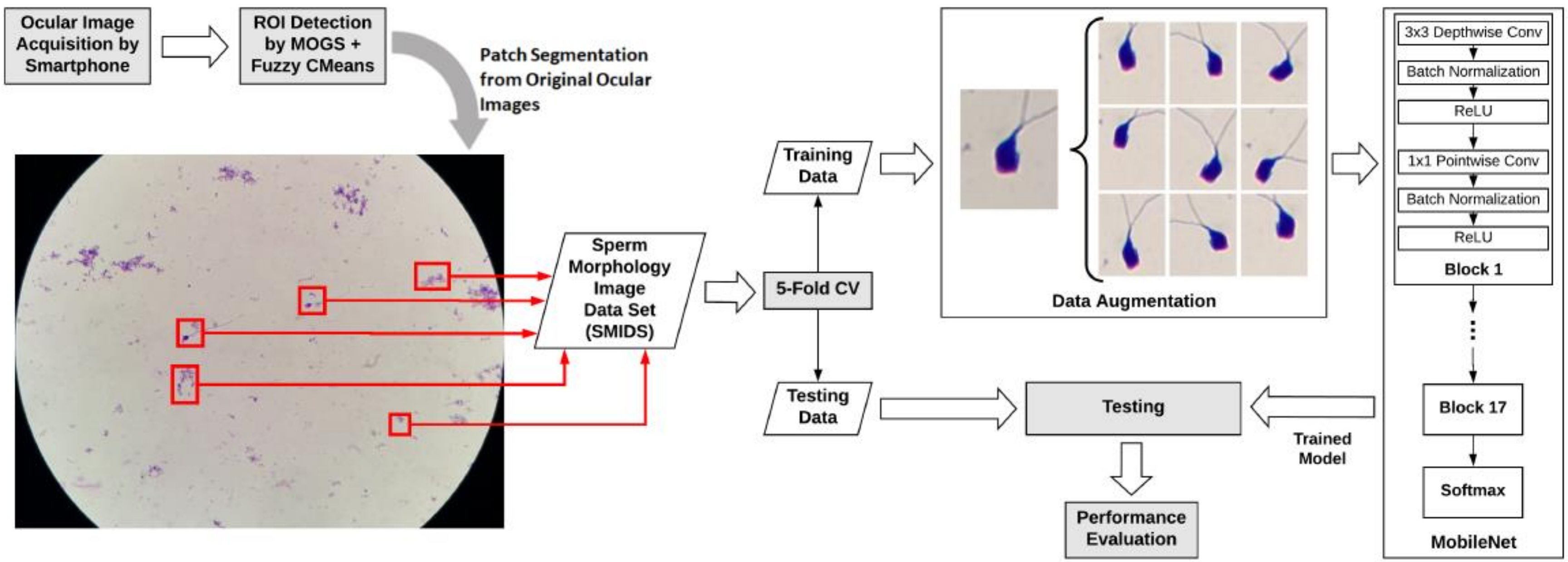
*Hamza Osman Ilhan, Görkem Serbes, Nizamettin Aydin, Automatic directional masking technique for better sperm morphology segmentation and classification analysis, Electronics Letters, 55.5 (2019): 256-258.

Sperm Morphology Analysis



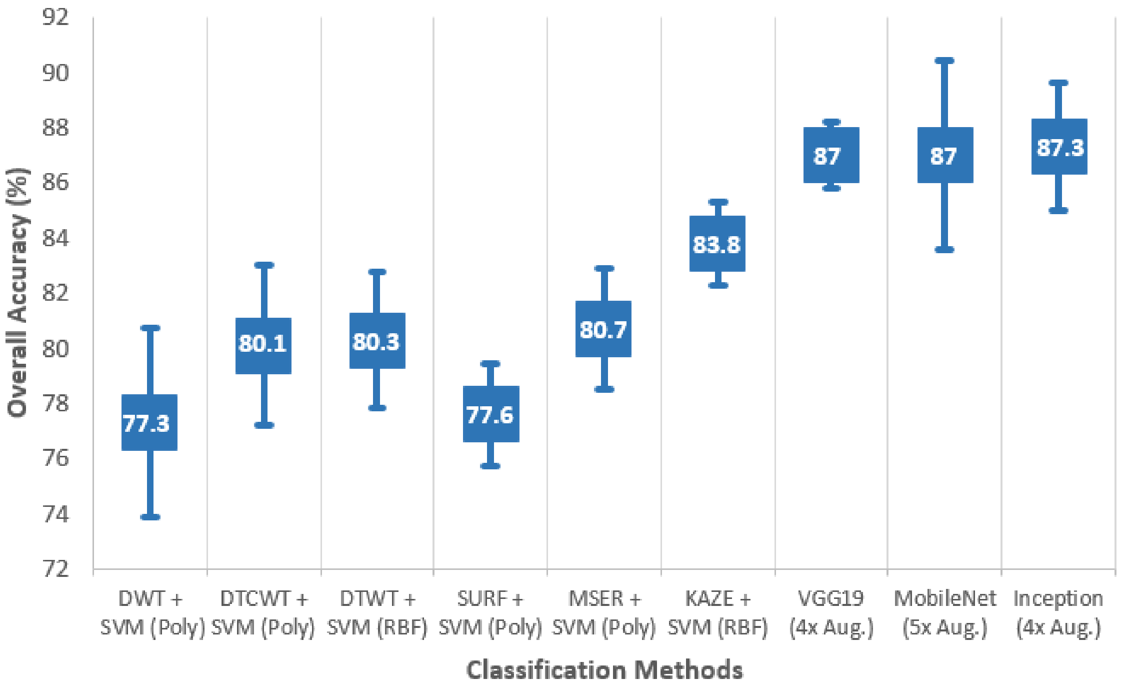
*Hamza Osman Ilhan, Görkem Serbes, Nizamettin Aydin, A Fully Automated Hybrid Human Sperm Detection and Classification System based on Mobile-Net and the Performance Comparison with Conventional Methods, Medical & Biological Engineering & Computing, (Under Review)

Sperm Morphology Analysis



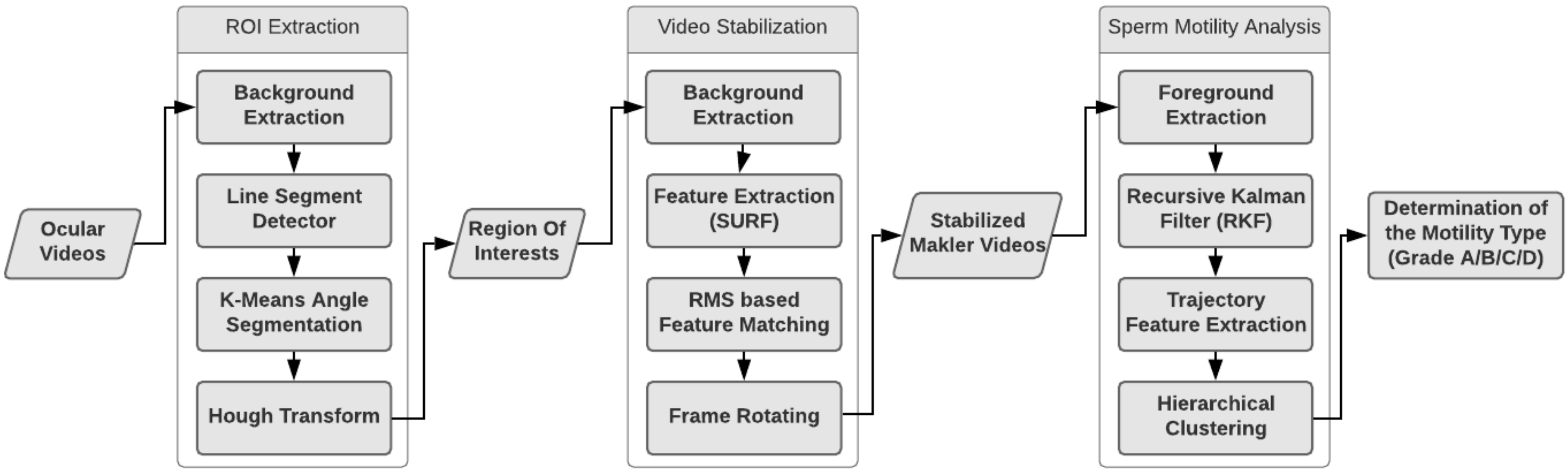
*Hamza Osman Ilhan, Görkem Serbes, Nizamettin Aydin, A Fully Automated Hybrid Human Sperm Detection and Classification System based on Mobile-Net and the Performance Comparison with Conventional Methods, Medical & Biological Engineering & Computing, (Under Review)

Sperm Morphology Analysis



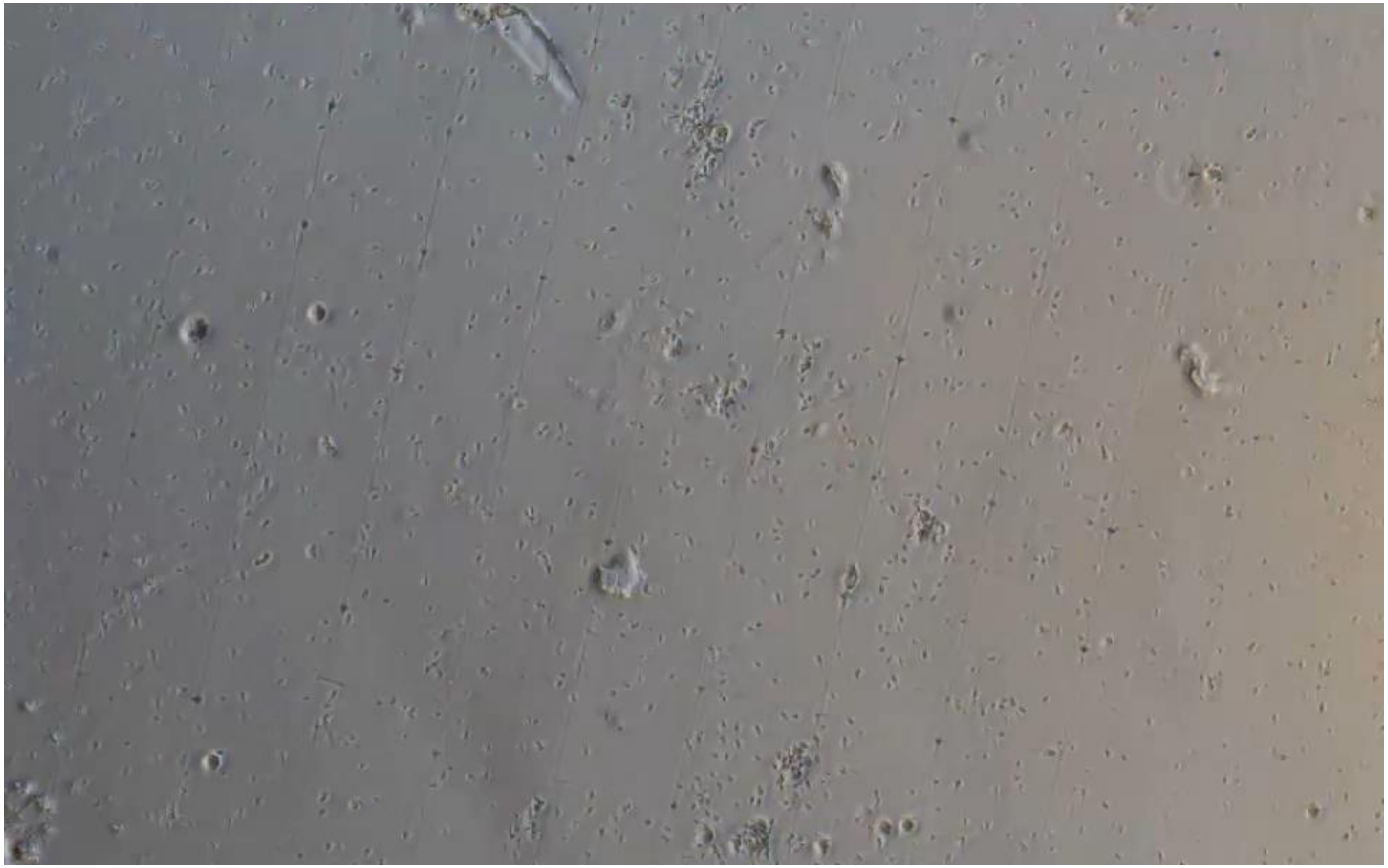
*Hamza Osman Ilhan, Görkem Serbes, Nizamettin Aydin, A Fully Automated Hybrid Human Sperm Detection and Classification System based on Mobile-Net and the Performance Comparison with Conventional Methods, Medical & Biological Engineering & Computing, (Under Review)

Sperm Motility Analysis

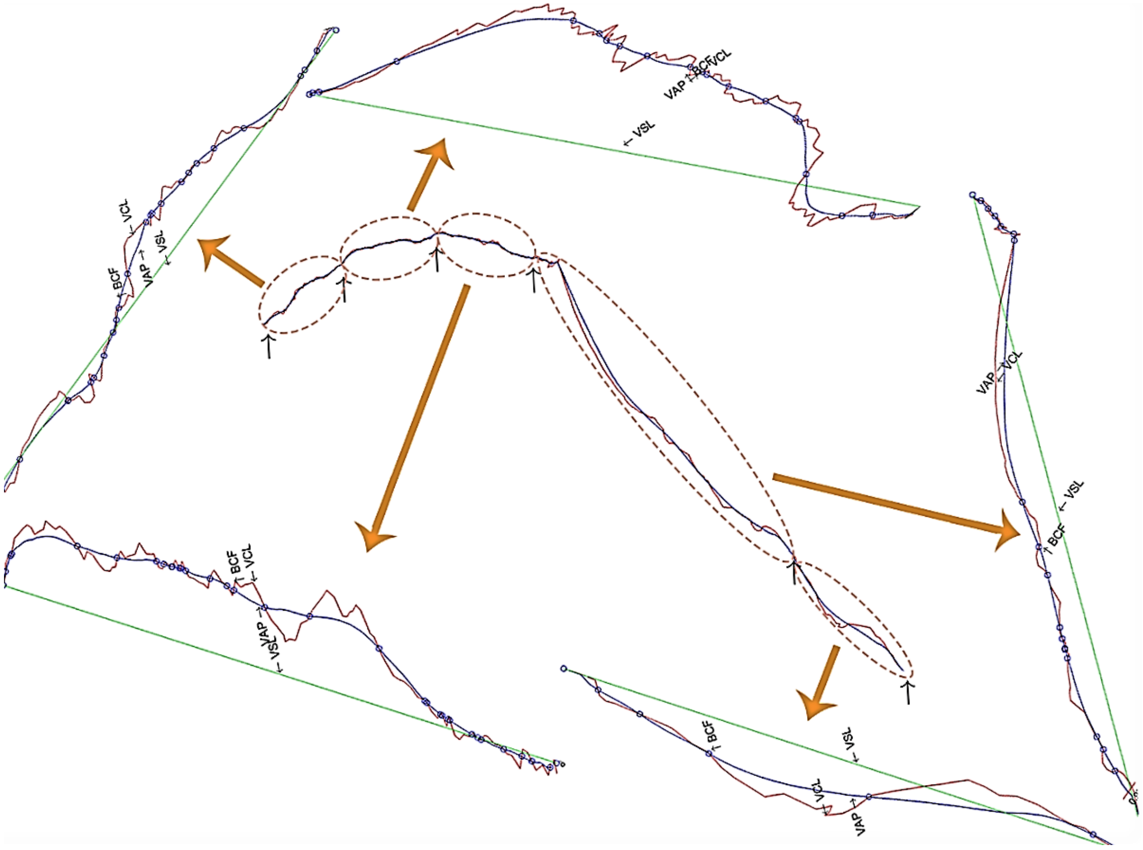


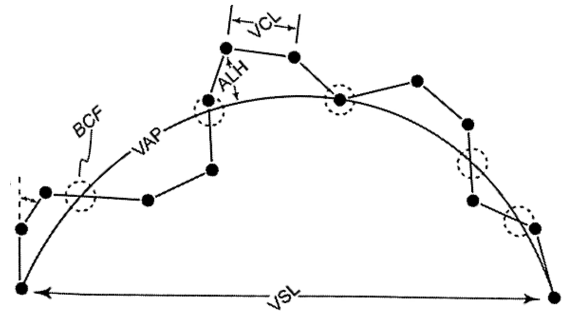
Sperm Motility Analysis

Sperm Motility Analysis



Sperm Motility Analysis



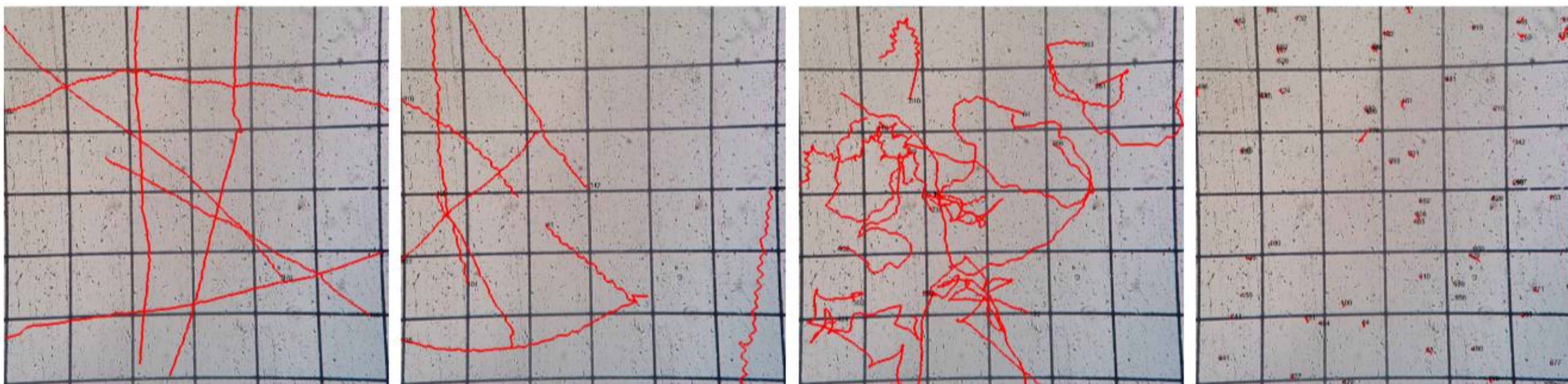


Trajectory Feature Extraction

- VSL
- VCL
- ALH
- LIN
- STR
- BCF
- VAP

Sperm Motility Analysis

	VCL	VSL	VAP	STR	LIN	WOB	BCF
Cluster 1 (Grade A)	35.3	29.2	**32.7**	89.2	**82.7**	92.6	11
Reference Grade A	-	-	**>25**	-	**>75**	-	-
Cluster 2 (Grade B)	21.8	17.4	**19.4**	89.3	**79.5**	88.6	14
Reference Grade B	-	-	**<25**	-	**>75**	-	-
Cluster 3 (Grade C)	11.8	4.9	8.5	57.2	**40.8**	67.9	17
Reference Grade C	-	-	-	-	**<75**	-	-
Cluster 4 (Grade D)	3.1	0.8	**1.4**	42.3	17.9	39.1	35
Reference Grade D	-	-	**<5**	-	-	-	-



a *b* *c* *d*

Figure 11: Motility Feature clustering over a subject a) Grade A Fast progressive, b) Grade B Progressive, c) Grade C Non-progressive, d) Grade D Stable

Example - 5

- Music Genre Classification and Recommendation by Using Machine Learning Techniques
  - Music genre prediction is the one of the topics that digital music processing is interested in.
    - In this study:
      - Acoustic features of music have been extracted by using digital signal processing techniques.
      - Music genre classification and music recommendations have been made by using machine learning methods.
        - **without** considering **user's music profile** or **collaborative filtering**.

- Convolutional Neural Networks were used for genre classification and music recommendation.
- The features to be extracted from music has determined as
  - Zero Crossing Rate,
  - Spectral Centroid,
  - Spectral Constrast,
  - Spectral Bandwitdth,
  - Spectral Rollof,
  - Mel-Frequency Cepstral Coefficients (MFCC).
- To compare performance results, all classification and recommendation algorithms has been applied on the GTZAN dataset.

DATASET

- GTZAN dataset is used due to its popularity in music signal processing.
  - Dataset firstly proposed by G. Tzanetakis in IEEE Transactions on Audio and Speech Processing 2002.
- Dataset contains **1000 audio tracks** which includes **100 tracks** for each **10 genres**.
  - Genres; **Blues**, **Classical**, **Country**, **Disco**, **Hiphop**, **Jazz**, **Metal**, **Pop**, **Reggae**, **Rock**.
  - Each tracks are **30 seconds** long, with **22050Hz** sampling frequency and **16-bits**.

FEATURE EXTRACTION

- 7 feature extraction techniques used:
  - Zero Crossing Rate, Spectral Centroid, Spectral Contrast, Spectral Bandwidth, Spectral Rollof, MFCC, MFCC Derivative.

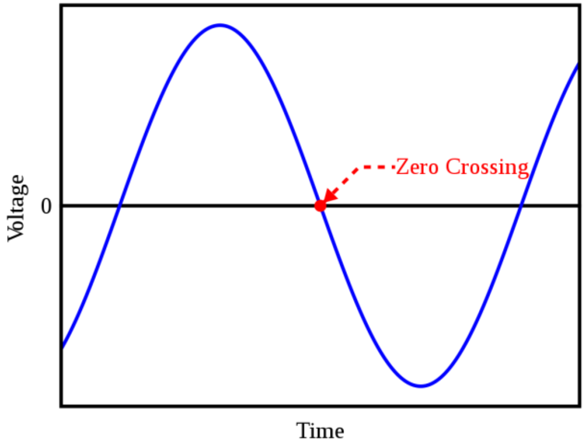
Feature Extraction Method	Descriptive Statistics	# of Features
ZCR	arithmetic mean arithmetic median standard deviation	3
Spectral Centroid		3
Spectral Bandwidth		3
Spectral Contrast		3
Spectral Rollof		3
MFCC (13 coeff)		39
MFCC Derivative (13 coeff)		39
	TOTAL	93

ZERO CROSSING RATE

- Ratio of sign changes
  - Negative to positive, positive to negative
    - More changes: High frequency
    - Less changes: Low frequency

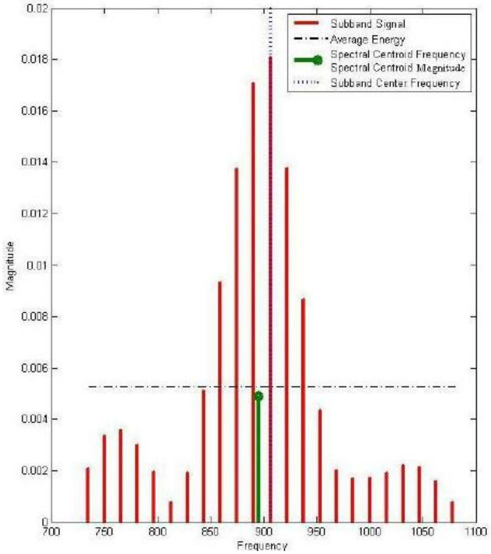
$$f(x) = \frac{1}{T-1}\sum_{i=1}^{T-1} f(s_t s_{t-1} < 0)$$

- **s:** signal
- **T:** signal length



SPECTRAL CENTROID

- Brightness of a sound
  - Amount of high frequency
  - Strong distinction
- Loudness
- Center of weight of the spectrum



$$Centroid, \mu = \frac{\sum_{i=1}^{N} f_i \times m_i}{\sum_{i=1}^{N} m_i}$$

- $\boldsymbol{f_i}$: center frequency of that bin
- $\boldsymbol{m_i}$: magnitude of bin number i

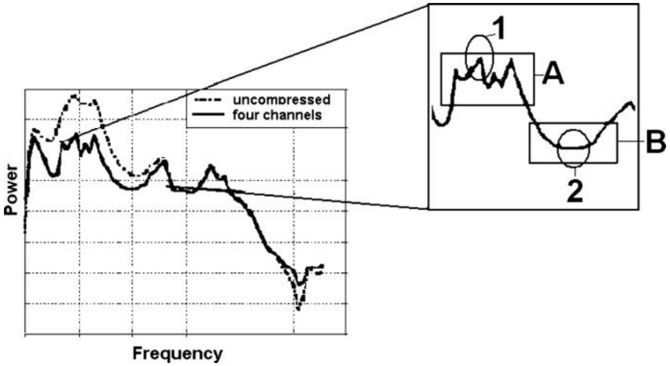
SPECTRAL CONTRAST

- Peaks and valleys in spectrum
  - Difference between them: decibel difference
    - Strong spectral peaks: harmonic components
    - Valleys: noise

$$Peak_k = \log\left\{\frac{1}{\alpha \times N}\sum_{i=1}^{\alpha \times N} x'_{k,i}\right\} \quad Valley_k = \log\left\{\frac{1}{\alpha \times N}\sum_{i=1}^{\alpha \times N} x'_{k,N-i+1}\right\}$$

$$Difference = Peak_k - Valley_k$$

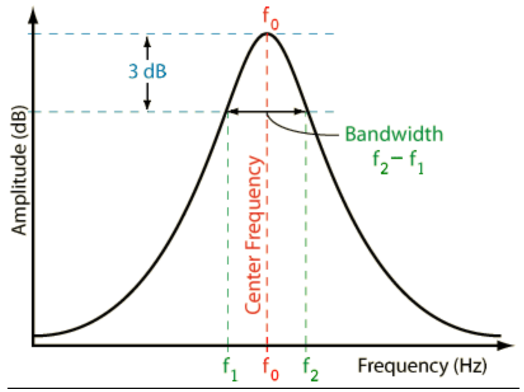
- **k:** sub-band number of the FFT vector $\{x_{k,1}, x_{k,2}, \dots, x_{k,N}\}$
- **N:** total number in k-th sub-band
- **α:** constant



SPECTRAL BANDWIDTH

- Frequency range in the frame
- Weighted average amplitude difference between frequency magnitude

$$\left(\sum_k S(k) \times (f(k) - f_c)^p\right)^{\frac{1}{p}}$$

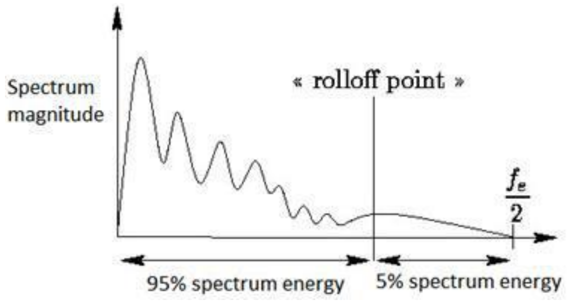


- $\boldsymbol{S(k)}$ **:** spectral magnitude at frequency bin k
- $\boldsymbol{f(k)}$**:** frequency at bin k
- $\boldsymbol{f(c)}$**:** spectral centroid
- $\boldsymbol{p = 2}$**:** weighted standard deviation

SPECTRAL ROLLOF

- Distinguish between rapid and slow music
  - Frequency value at which a magnitude distribution below a certain percentage value concentrates

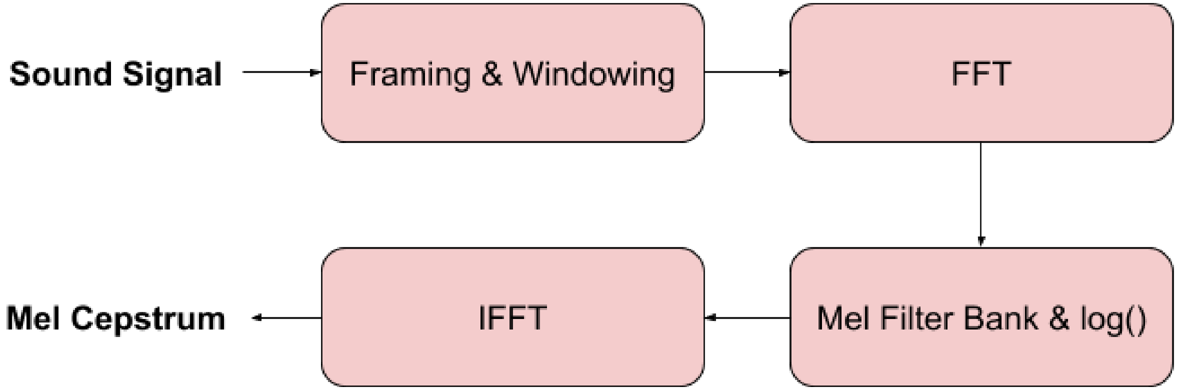
$$\sum_{n=1}^{R_i} M_i[n] = 0.85 \times \sum_{n=1}^{N} M_i[n]$$



- $\boldsymbol{M_i[n]}$ **:** magnitude of the Fourier transform at frame t and frequency bin n

MEL FREQUENCY CEPSTRUM COEFFICIENTS (MFCC)

- Most popular feature extraction technique
  - Purpose: Cepstral coefficients ➔ Human hearing system
  - Cepstral coefficients:
    - Normally: Linear scale
    - In MFCC:
      - Below 1 kHz: Linear scale
      - Upper: Logaritmic scale



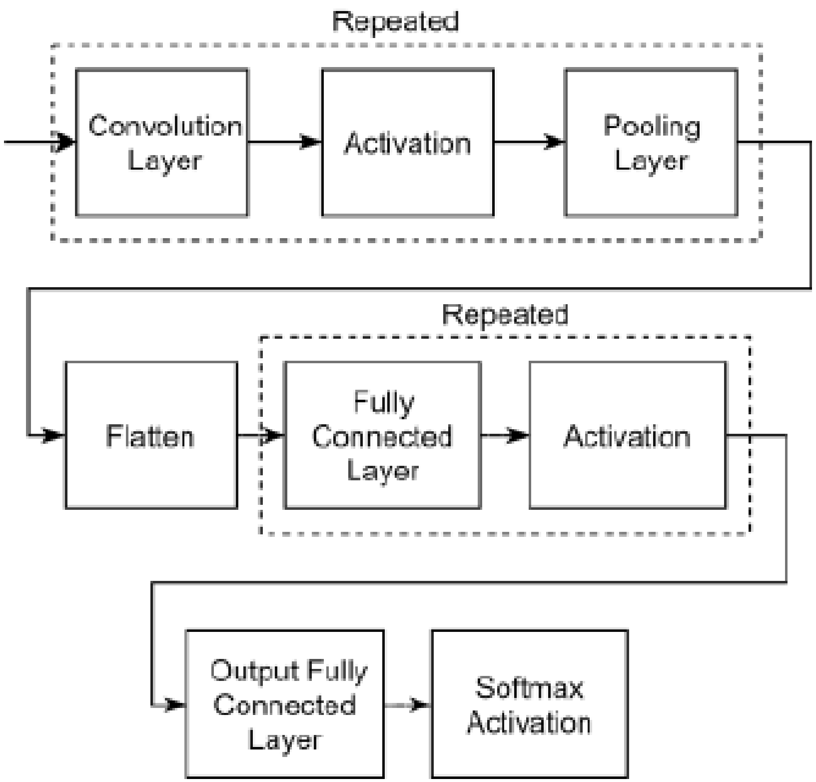
MACHINE LEARNING

- 5 classifier used:
  - K-Nearest Neighbor (K-NN),
  - Naive Bayes (NB),
  - Decision Tree (DT),
  - Support Vector Machine (SVM),
  - Random Forest (RF).

DEEP LEARNING

- Convolutional Neural Networks used to classify items that contain spatial neighborhood.
- In this study, array of randomly created filters are used and they are tweaked to better describe the data.
- Even though, they are used to classify images but one dimensional filters can be utilized to classify audio.

DEEP LEARNING



1- 2D Convolution Layer, 5x5 sized 32 filters, LeakyReLU activation function
2- Max Pooling Layer
3- 2D Convolution Layer, 5x5 sized 32 filters, LeakyReLU activation function
4- Max Pooling Layer
5- 2D Convolution Layer, 5x5 sized 32 filters, LeakyReLU activation function
6- Max Pooling Layer
7- 2D Convolution Layer, 5x5 sized 32 filters, LeakyReLU activation function
8- Average Pooling Layer
9- Flatten Layer
10- Dense Layer, 256 nodes, LeakyReLU activation function
11- Dense Layer, 128 nodes, LeakyReLU activation function
12- Dense Layer, 64 nodes, LeakyReLU activation function
13- Dense Layer, 10 nodes, Softmax activation function, as output layer

DEEP LEARNING

- The inputs of network are as follows:
  - Raw Audio:
    - Directy fed to the network. 1D convolutional layers are used.
  - Short Time Fourier Transform (STFT):
    - Applied to audio before feding. This transforms 1D time series into 2D frequency domain.
  - MFCC:
    - Mel Frequency Cepstrum is a representation of the short-term power spectrum of a sound, based on mel scaled spectrogram. MFC Coefficients make up the MFC.

RESULTS

Algorithm	hann		barthann	
	1024	4096	1024	4096
KNN	61.50%	62.99%	61.70%	62.50%
RF	61.90%	63.69%	62.80%	65.69%
NB	55.30%	56.90%	55.09%	56.20%
DT	48.50%	55.30%	50.90%	55.00%
SVM	**72.60%**	**72.70%**	**72.30%**	**72.39%**

Classification Results by Using All Features

- **Window type**: Hanning(hann) and Bartlett-Hann(barthann)
- **Window size:** 1024 and 4096
- **K**: 3 for KNN
- **Kernel**: Linear for SVM

RESULTS

Genre	First 5 Songs	First 10 Songs
Blues	60%	48%
Classical	88%	90%
Country	40%	50%
Disco	48%	40%
Hip-hop	72%	60%
Jazz	52%	42%
Metal	84%	76%
Pop	60%	50%
Reggae	24%	26%
Rock	60%	50%

Recommendation Results by Using Conventional Features

RESULTS

Data Type	Accuracy	Recall	Precision	F-Measure
Raw Music	15.00 %	19.00%	15.00%	13.00%
STFT	**66.00 %**	**65.00 %**	**69.00 %**	**65.00 %**
MFCC	63.00 %	63.00 %	64.00 %	62.00 %

CNN Classification Results

RESULTS

Genre	First 5 Songs	First 10 Songs
Blues	63.00%	49.70%
Classical	90.80%	87.70%
Country	56.80%	49.40%
Disco	53.60%	45.90%
Hip-hop	64.20%	57.30%
Jazz	65.00%	52.70%
Metal	79.40%	75.10%
Pop	78.40%	73.80%
Reggae	57.20%	48.90%
Rock	49.80%	42.20%

CNN Recommendation Results

RESULTS

- SVM achieved better classification results than other methods.
- Changing the window size and window type caused very small performance changes.
- MFCC has better effect than the other methods.
- Using deep learning method showed that there is no considerable performance change on music genre classification.
- SVM has achieved higher success than the CNN algorithm.
- For some genres of music like Classical, the music recommendation is highly successful, while in some species the performance falls.

REFERENCES

[1] A. Tzanetakis, G. and Cook, P. "Musical genre classification of audio signal", IEEE Transactions on Speech and Audio Processing, Vol. 10, No. 3, pp. 293-302, July 2002.

[2] Holzapfel, A. and Stylianou Y. "Musical genre classification using nonnegative matrix factorization-based features", IEEE Transactions on Audio, Speech, and Language Processing, Vol. 16, No. 2, pp. 424-434, 2008.

[3] Benetos, E. and Kotropoulos C. "A tensor-based approach forautomatic music genre classification", Proceedings of the European Signal Processing Conference, Lausanne, Switzerland, 2008

[4] Lidy, T., Rauber, A., Pertusa, A. and Inesta, J. "Combining audio and symbolic descriptors for music classification from audio", Music Information Retrieval Information Exchange (MIREX), 2007.

[5] Bergstra, J., Casagrande, N., Erhan, D., Eck, D. and Kegl B. "Aggregate features and AdaBoost for music classification", Machine Learning, Vol. 65, No. 2-3, pp. 473-484, 2006.

[6] A. Karatana and O. Yildiz, "Music genre classification with machine learning techniques," in Signal Processing and Communications Applications Conference (SIU), 2017 25th, IEEE, 2017, pp. 1–4.

[7] Han, J., Kamber, M., & Pei, J. (2011). Data mining: concepts and techniques: concepts and techniques. Elsevier.

[8] Witten, I. H., & Frank, E. (2005). Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann.

[9] Cortez P., In press. RMiner: Data Mining with Neural Networks and Support Vector Machines using R. In R. Rajesh (Ed.), Introduction to Advanced Scientific Softwares and Toolboxes.

[10] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5– 32, Oct. 2001.

[11] G. Gwardys and D. Grzywczak, "Deep image features in music information retrieval," Intl Journal of Electronics and Telecomunications, vol. 60, no. 4, pp. 321–361, 2014.

[12] https://librosa.github.io/librosa/ (visited on 15/06/2018)

[13] https://keras.io (visited on 15/06/2018)

[14] S. Dieleman. (2014). Recommending music on spotify with deep learning.[Online].Available: http://benanne.github.io/2014/08/05/spotifycnns.html (visited on 03/31/2018

[15] Ahmet Elbir, Hamza Osman Ilhan, Gorkem Serbes, Nizamettin Aydin. "Short Time Fourier Transform based music Genre classification" , 2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT), 2018