International Conference on Recent Trends in Engineering & Technology ICRTET-2013

Prof. Dr. Srikanta Patnaik

Follow this and additional works at: https://www.interscience.in/conf_proc_volumes

Part of the Engineering Commons
The new economic millennium is surprisingly rushing with Innovation and Technology. There is a sharper focus on deriving value from the widespread global integration in almost all spheres of social, economic, political and technological subsystems. In the quest of making this earth a better place to live we have to make a strong hold upon sustainable energy source. Sustainable energy sources include all renewable energy sources, such as hydroelectricity, solar energy, wind energy, wave power, geothermal energy, bio-energy, and tidal power. It usually also includes technologies designed to improve energy efficiency. Energy efficiency and renewable energy are said to be the twin pillars of sustainable energy. Renewable energy technologies are essential contributors to sustainable energy as they generally contribute to world energy security, reducing dependence on fossil fuel resources, and providing opportunities for mitigating greenhouse gases. Although the discipline like electrical engineering has narrated academic maturity in the last decades, but the limitations of the non-renewable energy sources, turbulence and disturbances in the energy propagation cascades various insightfulness and stimulation in post classical electrical era. Evidence shows that there are phenomenal supplements in power generation and control after the introduction of Energy Management System (EMS) supported by Supervisory Control and Data Acquisition (SCADA). As there is increasing focus on strengthening the capacity of the power houses with the existing resources or constraints some new dimensions like FACTS, Optimal System Generation, High Voltage DC transmission system, Power Generation Control, Soft Computing, Compensation of transmission line, Protection scheme of generator, Loss calculation, economics of generation, fault analysis in power systems are emerging. Since the world is suffering with water, food, and energy crisis, energy consumption has social relevancy.

Let me highlight some of the recent developments in Electronics discipline. The new integrated devices did not find a ready market. Users were concerned because the individual transistors, resistors, and other electronic circuit components could not be tested individually to ensure their reliability. Also, early integrated circuits were expensive, and they impinged on the turf that traditionally belonged to the circuit designers at the customer's company. Again, Bob Noyce made a seminal contribution. He offered to sell the complete circuits for less than the customer could purchase individual components to build them. (It was also significantly less than it was costing us to build them!) This step opened the market and helped develop the manufacturing volumes necessary to reduce manufacturing costs to competitive levels. To this day the cost reductions resulting from economies of scale and newer high-density technology are passed on to the user—often before they are actually realized by the circuit manufacturer. As a result, we all know that the high-performance electronic gadget of today will be replaced with one of higher performance and lower cost tomorrow.

The integrated circuit completely changed the economics of electronics. Initially we looked forward to the time when an individual transistor might sell for a dollar. Today that dollar can buy tens of millions of transistors as part of a complex circuit. This cost reduction has made the technology ubiquitous—nearly any application that processes information today can be done most economically electronically. No other technology that I can identify has undergone such a dramatic decrease in cost, let alone the improved performance that comes from making things smaller and smaller. The technology has advanced so fast that I am amazed we can design and manufacture the products in common use today. It is a classic case of lifting ourselves up by our bootstraps—only with today's increasingly powerful computers can we design tomorrow's chips.

The conference is designed to stimulate the young minds including Research Scholars, Academicians, and Practitioners to contribute their ideas, thoughts and nobility in these two integrated disciplines. Even a fraction of active participation deeply influences the magnanimity of this international event. I must acknowledge your response to this conference. I ought to convey that this conference is only a little step towards knowledge, network and relationship.

I express best wishes to all the paper presenters. I extend my heart full thanks to the reviewers, editorial board members, programme committee members of the conference. If situations prevail in favor we will take the glory of organizing the second conference of this kind during this period next year.

Convenor :-

Mr. Bikash Chandra Rout
EDGE DETECTION OF MRI SCAN NECK IMAGE USING CANNY EDGE DETECTION ALGORITHM BASED ON DENSITY OF EDGE LENGTH

PREETHA J1 & SELVARAJAN S2

1Department of Computer Science and Engineering
Muthayammal Engineering College, Rasipuram – 637 408, Tamilnadu, India
2 Muthayammal Technical Campus, Rasipuram – 637 408, Tamilnadu, India.

Abstract- Neck is robot part for human. The robotic platform of the neck bone was surrounded by muscles like steel spring the neck has a significant amount of movement and supports the weight of the head, but it is less protected than the rest of the neck, the neck can be vulnerable to injury and disorders that produce pain and restrict motion. In this work Density of Edge Length based Canny Edge Detection Algorithm has been proposed to pre processing of boundary detection of the MRI Scan Neck image. To find the correct boundary in noisy image of neck is still a difficult one. The proposed Canny Edge Detection algorithm has been used to detect the boundaries of neck image from the noisy image. The performance of proposed technique has been verified and validated with the standard medical values. The results show that the proposed technique performs well and produced very near to the optimal solution. This method is robust for all kinds of noisy images.

Keywords - MRI Scan, Neck image, Edge Length and Canny Edge Detection Algorithm

I. INTRODUCTION

Manual Material Handling (MMH), especially lifting, poses a risk to many and considered the prime cause of back pain and various other joint impairments. This in turn leads to increased worker compensation and loss of productive man-hours. Approximately one third of all jobs in industry involve MMH. Low back pain is one of the most prevalent and costly work related injuries. The study of the lifting posture, the amounts of weight the man can safely lift are the areas concentrated on by Researchers. Of all Manual Material Handling (MMH) activities, lifting is considered to be a major cause for neck pain, low back pain and neck injuries. Medical image processing is a field of science due to its advances technology and software breakthroughs. It plays a vital role in disease diagnosis and improved in medical area with patient care and helps medical practitioners to make decision in regard to the type of treatment. Various state-of-the-art equipments are available for human organs in digital form. Examples of such devices include X-Ray-based devices, Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Ultrasound (US), Positron Emission Tomography (PET) and Single Photon Emission Computed Tomography (SPECT). In which, X-Ray is one the oldest and frequently used devices, as they are non-invasive, painless and economical. A x-ray makes images. of any bone in the body, including the hand, wrist, arm, elbow, shoulder, foot, ankle, leg, knee, thigh, hip, pelvis or neck.

![Fig. 1. Human Posture Working with VDT](image)

Most cervical problems are due to degenerative changes that occur in the s and joints of the neck. Degenerative changes that affect the structures of the neck can cause the neck canal to become too narrow, a condition called neck stenosis. This may lead to pressure on the neck cord. Bone spurs that stick into the neck canal take up space, making the neck canal smaller. They can press against the neck cord or nerve roots. Pressure on the neck cord from neck stenosis can cause symptoms of myelopathy. Myelopathy may impair normal walking, hand and finger use, and bowel and bladder function. Doctors take these symptoms very seriously because severe myelopathy that is not treated may lead to permanent nerve or neck cord damage. Pressure on nearby nerve roots can cause...
radiculopathy and may produce pain, weakness, or sensory changes in the area supplied by nerves that go from the cervical neck to the shoulder, arm, or hand. The disk acts as a shock absorber between the bones in the neck. In cervical disk degeneration the normal gelatin-like centre of the disk degenerates and the space between the vertebrae narrows. As the disk space narrows, added stress is applied to the joints of the neck causing further wear and degenerative disease. The cervical disk may also protrude and put pressure on the neck cord or nerve roots when the rim of the disk weakens. This is known as a herniated cervical disk.

The neck is so flexible and it supports the head, it is extremely vulnerable to injury. Motor vehicle or diving accidents, contact sports, and falls may result in neck injury. The regular use of safety belts in motor vehicles can help to prevent or minimize neck injury. A "rear end" automobile collision may result in hyperextension, a backward motion of the neck beyond normal limits, or hyperflexion, a forward motion of the neck beyond normal limits. The most common neck injuries involve the soft tissues: the muscles and ligaments. Severe neck injuries with a fracture or dislocation of the neck may damage the neck cord and cause paralysis. The main focus of this work is to automatically detect the reason for neck pain from plain diagnostic x-rays using a series of sequential steps. Fig.1 shows the schematic diagram of Human neck.

A. Bone Spurs in the Back and Neck

Bone spurs, also called osteophytes, are enlargements of your body’s normal bone structure. As these enlargements progress, they may protrude into surrounding tissues, sometimes causing pain and other symptoms. Bone spurs can occur on virtually any bone, including on your vertebrae, but contrary to their name, they do not actually create a point. Rather, they are smooth structures that your body creates to repair itself after bones are exposed to pressure, stress, or rubbing over time. Bones conform to any pressure applied to them, and through the years, tendons (which hold muscle to bone) and ligaments (which hold bone to bone) can start to pull the bone from where it should be, stimulating the creation of bone spurs. As these protrusions grow and form, they sometimes impinge on a nerve, causing pain and debilitating symptoms. This figure is Bony Spurs in neck, colored X-ray. Bony spurs (osteophytes) grow from the vertebrae as a normal part of the aging process. They stabilize the neck when the s of cartilage between the vertebrae shrink and become less flexible.

- Types of Bone Spurs
  - Neck Bone Spurs
  - Neck Bone Spurs
  - Cervical Bone Spurs
  - Lumbar Bone Spurs

B. Neck Bones

- There are 7 vertebrae (bones) in the neck.
- Together they form the upper-most section of the vertebral column which is known as the "cervical neck".
- They are labelled C1 to C7 with C1 at the top and C7 furthest from the head and adjoining the first vertebra of the next section of the neck.
- Only two of the vertebrae of the cervical neck also have individual names. They are C1 which is called the "Atlas" bone, and C2 which is called the "Axis" bone. The atlas and axis bones are labelled on the diagram (above-right).

Table 1. Shows the types of neck pains an its causes and Figure 3 shows the description of the neck image. Figure 4 Shows the MRI Scan Human neck image.

Table 1. Types of neck pain and its causes

<table>
<thead>
<tr>
<th>S. No</th>
<th>Neck Pain cause / syndrome</th>
<th>Description / Explanation(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Postural neck pain</td>
<td>Postural neck pain refers to pain in the neck and/or shoulders caused or exacerbated by postural habits such as holding the head/neck/shoulders in fixed protracted positions for long periods of time</td>
</tr>
<tr>
<td>2</td>
<td>Acute neck pain (unknown cause)</td>
<td>If a sudden movement of the neck results in severe neck pain possibly accompanied by arm pain, muscle spasm and/or restricted movement of the neck then the problem resolves itself without intervention or identification of the structural cause</td>
</tr>
<tr>
<td>3</td>
<td>Cervical spondylosis</td>
<td>Cervical spondylosis results from on-going 'wear and tear' of the cervical vertebrae and the intervertebral s that separate</td>
</tr>
</tbody>
</table>
Cervical myelopathy is due to pressure on the neck cord leading to dysfunction of the nerves below the area of pressure. It is therefore a condition of the nerves.

In general this can result in joints becoming painful, swollen, and stiff. Rheumatoid arthritis in the neck (cervical neck) often involves the atlantoaxial joint which is the articulation between C1 (the "atlas" bone) and C2 (the "axis" bone).

In general, thoracic outlet syndrome can affects nerves, result in pressure on blood vessels - affecting circulation, and cause pain, tingling, swelling and/or weakness.

"Tumours originating in the cervical neck (neck bones) are rare. When tumours are found in this area they are often found to have originated elsewhere in the body and are therefore said to be "secondary deposits".

Osteitis is a general term for inflammation of bone. Osteitis in the neck area (called "cervical neck osteitis") is uncommon, but not unheard-of, in many developed countries.

The empirical quantities of the neck bone components used in shirazi-Adl’s study are shown in Table 2.

II. PROPOSED WORK SEQUENCE

The Block diagram of the proposed system of Image edge detection technique is shown in Fig.5. The different process sequence is involved in this system is given in below. The Original image is obtained from the MRI Scan image centre and then it will be incorporated by using canny Edge Detection algorithm based on density of edge length. The results have been verified and analysed.
III. CANNY EDGE DETECTION ALGORITHM

The Canny algorithm can be used as an optimal edge detector based on a set of criteria which include finding the most edges by minimizing the error rate, marking edges as closely as possible to the actual edges to maximize localization, and marking edges only once when a single edge exists for minimal response. According to Canny, the optimal filter that meets all three criteria above can be efficiently approximated using the first derivative of a Gaussian function.

\[
G_{\text{Canny}}(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{1}
\]

Step 1: Calculate the average magnitude

\[
M(1, 2) = \frac{1}{N^2} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \sqrt{M_{x,1,2}^2 + M_{y,1,2}^2} \tag{2}
\]

Step 2: Calculate the density of the edge length. The density of the edge length is calculated from

\[
L(1, 2) = \frac{C(1, 2)}{\text{max}(1, 2)} \tag{3}
\]

Where \(C(i,j)\) is the number of connected pixels at each position of pixel.

Step 3: Calculate the Initial position of map from summation of density of edge length and average magnitude.

\[
P(1, 2) = \frac{1}{2(M(1, 2) + L(1, 2))} \tag{4}
\]

Step 4: Calculate the thresholding of the initial position map. If

\[
P(1, 2) \geq T_{\text{max}} \tag{5}
\]

Then \(P(1, 2)\) is the initial position of the edge following. And then we obtained the initial position by setting \(T_{\text{max}}\) to 92% of the maximum value.

Fig.6. (a). MRI Scan Noisy Neck image, (b). Average Magnitude Image, (c). Density of the Edge Length, (d). Processing Sequence, (d). Initial Position map, (e). Final Thresholding of edge map

From the above figure 6(a) to 6(e) and the analysis we can able to predict the proper and suitable initial position.

A. Final Filtering

Using a statistical test and we are able to obtain the intensity of the neck image.
Statistical test is also known as Hypothesis test. The filtering process is ended with high resolution edge value. The Fig 7(a) to 7(e) shown that the final filtering process sequence with final filtered images. The output image is having high resolution of filtered surface. Fig. 8 Shows the Graphical noise signal in each stage.

Fig 7 (a). A Normal Neck image

Fig 7 (b). Initial filtered image

Fig 7 (c). Binary Edge Detection image

Fig 7 (d). Edge corrected image

Fig 7 (e). Final corrected edge filtered image

Fig 8 (a). A Normal Neck image Noise Signal

Fig 8 (b). Initial filtered image Noise Signal

Fig 8 (c). Binary Edge detection image Noise Signal
IV. RESULTS AND DISCUSSION

To further evaluate the efficiency of the proposed method in addition to the visual inspection, the proposed boundary detection method numerically using the Hausdorff distance and the probability of error in image segmentation. Where \( P(O) \) and \( P(B) \) are probabilities of objects and background in images. The objects surrounded by the contours obtained using the five snake models and the proposed method are compared with that manually drawn by skilled doctors from the Medical Hospital.

Table 3. Average Results of Probability of Error in Image Segmentation

<table>
<thead>
<tr>
<th>S.N</th>
<th>Image illustration</th>
<th>Canny Edge Segmentation (%)</th>
<th>Medical Standard Value (%)</th>
<th>Error Difference (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MRI Scan noisy neck image</td>
<td>8.52</td>
<td>8.51</td>
<td>+0.01</td>
</tr>
<tr>
<td>2</td>
<td>Average magnitude image</td>
<td>7.76</td>
<td>7.73</td>
<td>+0.03</td>
</tr>
<tr>
<td>3</td>
<td>Density of edge length image</td>
<td>7.01</td>
<td>7.00</td>
<td>- 0.01</td>
</tr>
<tr>
<td>4</td>
<td>Filtered MRI Scan image</td>
<td>6.43</td>
<td>6.44</td>
<td>- 0.01</td>
</tr>
</tbody>
</table>

From the above Table 3 shows the average results of probability of Error in Image segmentation of canny edge detection algorithm and Medical standard value and also predicts the error difference. Showing the results it shows the Error difference value is very minimal and also negligible. So the proposed techniques produced nearer to the standard value. Fig.9 Shows the comparative analysis of canny edge detection value and the Medical standard value which is collected from the standard Hospital.

V. CONCLUSION

The proposed technique for boundary detection is applied it to neck image. Our edge following technique incorporates a vector image model and the edge map information. The proposed technique was applied to detect the object boundaries of noisy MRI Scan neck image where the well defined edges were encountered. The opinions of the skilled doctors were used as the ground truths of interesting object of neck. Besides the visual inspection, the proposed method was verified and evaluated using the probability of error. The results of detecting the object boundaries in noisy images show that the proposed technique is very close to the standard value which was given by eminent doctors. We have successfully applied the edge following technique to detect the object boundaries of neck image. The proposed method can be applied not only for medical imaging, but can also be applied to any image processing problems.

REFERENCES

Edge Detection of MRI Scan Neck Image using Canny Edge Detection Algorithm Based on Density of Edge Length


PREDICTION AND ANALYSIS OF LAG OF ACCOMMODATION OF EYE USING RAF RULER AND TAGUCHI METHOD

VIJAYALAKSHMI R1 & SELVARAJAN S2

1Department of Computer Science and Engineering
Muthayammal Engineering College, Rasipuram – 637 408, Tamilnadu, India.
2Muthayammal Technical Campus,
Rasipuram – 637 408, Tamilnadu, India.

Abstract: One of the major problems faced by Visual Display Terminal (VDT) users is the eyestrain, a visual discomfort that occurs during their work. To optimise the various factors that causes the eyestrain of computer users for reduce the eyestrain. In this work Taguchi method used for optimise the factors viewing angle, viewing distance and exposure time are the major factors considered here for solving the problem. In this work Lag of accommodation is taken as response variable. For measuring the eyestrain the instrument RAF ruler is used. In this method Orthogonal arrays are used for design the experiments. And all the experimental investigations made and the data is stored for data mining.

Keywords: Viewing angle, Viewing distance, Exposure time, orthogonal array.

1. INTRODUCTION

The introduction of VDT in the workplace has led to numerous reports of health disorder from its use. The disorders are predominantly referable to the eye and musculoskeletal systems. Vision is the sole means by which most computers users obtain information to perform their work. Computer – related vision problems may be a significant impact on worker productivity and frequency of errors and may result in increased absences on the job and increased absences on the job and decreased personal enjoyment from computer use. The symptoms associated with computer use include: Eyestrain, fatigue, headache, blurred vision, dry or irritated eyes, neck and/or backache and diplopia. These problems usually can result from improper lighting, glare from screen, poor positioning to screen itself or copy material that is difficult to read. Figure 1. Shows the Human posture working with VDT environment.

These problems can be corrected by adjusting the physical and environmental setting where the VDT users work. The most common symptom reported by VDT operator has been “Eyestrain” more properly called “Visual fatigue”. Eyestrain is defined as a kind of visual discomfort, which occurs due to the prolonged attention of visual detail with reduced eye movement in a restricted visual field. Eyestrain can be measured in terms of Lag of accommodation. The aim of the work is to optimise the ergonomic factors which cause the visual discomfort to VDT users during their work. The work considered the ergonomic factors i) Viewing angle, ii) Viewing Distance of VDT users from the monitor and iii) Exposure Time.

2. MEASUREMENT OF FACTORS

Eyestrain due to prolonged of Visual Display Terminal. Objective of this work is to increase the productivity of the VDT user.

2.1 MEASUREMENT SYSTEM

i). Lag of accommodation as a response variable

ii). Measuring equipment is RAF ruler.
(Shown in Figure 2). Figure 3 shows the way to use the RAF ruler.
Measurement of the interpupillary distance is made by measuring the distance A between the two corneal images, or B between the edges of the pupils (if both pupils are of the same size), or C between the edges of the limbus. In normal use, the PD rule is typically placed along the horizontal diameter of the cornea.

3. FACTORS THAT INFLUENCE THE EYESTRAIN

3.1 FIXED FACTORS
Seating arrangement, Colour of the monitor, Glare, Type of task, Illumination level.

3.2 VARIABLE FACTORS
- Viewing Angle (A)
- Viewing Distance (B)
- Exposure Time (C)

4. LEVELS OF THE FACTOR
The level for each factor was selected as three.

Table 1 shows the factors and levels

<table>
<thead>
<tr>
<th>FACTORS</th>
<th>LEVEL 1</th>
<th>LEVEL 2</th>
<th>LEVEL 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>A in degree</td>
<td>-25</td>
<td>-20</td>
<td>-15</td>
</tr>
<tr>
<td>B in cm</td>
<td>60</td>
<td>70</td>
<td>80</td>
</tr>
<tr>
<td>C in min</td>
<td>45</td>
<td>60</td>
<td>75</td>
</tr>
</tbody>
</table>

‘-’ sign indicates the angle formed below the eyeline.

4.1 ASSIGNMENT OF FACTORS AND INTERACTIONS
Table 2 shows the factors and interactions.

<table>
<thead>
<tr>
<th>Factors &amp; Interactions</th>
<th>Degrees of freedom</th>
<th>Column No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

4.2 RESOLUTION OF EXPERIMENTAL ARRAYS
Table 3 shows the factors and levels

<table>
<thead>
<tr>
<th>No. of Trials</th>
<th>No. of Levels</th>
<th>No. of factors</th>
<th>Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>27</td>
<td>3</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

5. EQUIPMENT USED
- Adjustable Chair
- Adjustable Table
- Platform
- Vertical Scale
- Eyeline Indicator
- RAF Ruler

5.1 VIEWING ANGLE MEASUREMENT
\[ h = D \tan \theta \]
Where,
- \( H \) = Equivalent height to be maintained for particular \( \theta \)
- \( D \) = Viewing Distance, cm
- \( \theta \) = Viewing Angle, deg.

5.2. EQUIVALENT HEIGHT FOR THE VIEWING ANGLES
Table 4 shows the angles and height

<table>
<thead>
<tr>
<th>D (cm)</th>
<th>( \theta ) (deg)</th>
<th>H (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>-15</td>
<td>16.0</td>
</tr>
<tr>
<td></td>
<td>-20</td>
<td>21.8</td>
</tr>
<tr>
<td></td>
<td>-25</td>
<td>28.0</td>
</tr>
<tr>
<td>70</td>
<td>-15</td>
<td>18.7</td>
</tr>
<tr>
<td></td>
<td>-20</td>
<td>25.5</td>
</tr>
<tr>
<td></td>
<td>-25</td>
<td>32.6</td>
</tr>
<tr>
<td>80</td>
<td>-15</td>
<td>21.4</td>
</tr>
<tr>
<td></td>
<td>-20</td>
<td>29.1</td>
</tr>
<tr>
<td></td>
<td>-25</td>
<td>37.3</td>
</tr>
</tbody>
</table>

6.0 EXPERIMENTAL RESULTS FOR LAG OF ACCOMMODATION
(1.27 Orthogonal Array)
Table 5 shows the measurement of Lag of Accommodation (RAF Ruler measurement) in mm.
7.0 CALCULATION FOR LAG OF ACCOMMODATION:

Total variation,

\[
SS_T = SS_A + SS_B + SS_C + SS_{AB} + SS_{AC} + SS_{BC} + SS_{ABC} + SS_e
\]

1. \[
F_1 = \frac{\text{Sum of Squares}}{\text{DOF}} \text{ Variance} \text{ F Ratio}
\]

\[
A = 209.4, 2, 104.7, 152.8^* \\
B = 55.15, 2, 27.57, 40.2^* \\
C = 128.5, 2, 64.24, 93.8^* \\
AXB = 26.07, 4, 6.519, 9.51^* \\
AXC = 1.407, 4, 0.352, 0.51 \\
BXC = 3.296, 4, 0.824, 1.2 \\
AXBXC = 5.148, 8, 0.644, 0.94
\]

* Indicates significant for 95% confidence
Prediction And Analysis of Lag of Accommodation of Eye using Raf Ruler and Taguchi Method

<table>
<thead>
<tr>
<th>Factors &amp; Interactions</th>
<th>Sum Of Squares</th>
<th>DOF</th>
<th>Variance SS</th>
<th>F Ratio</th>
<th>SS1</th>
<th>% Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>209.4</td>
<td>2</td>
<td>104.7</td>
<td>158.8</td>
<td>208.1</td>
<td>46.5</td>
</tr>
<tr>
<td>B</td>
<td>55.15</td>
<td>2</td>
<td>27.57</td>
<td>41.8</td>
<td>53.83</td>
<td>12.03</td>
</tr>
<tr>
<td>C</td>
<td>128.5</td>
<td>2</td>
<td>64.24</td>
<td>97.4</td>
<td>127.2</td>
<td>28.42</td>
</tr>
<tr>
<td>AXB</td>
<td>26.07</td>
<td>4</td>
<td>6.519</td>
<td>9.89</td>
<td>23.44</td>
<td>5.238</td>
</tr>
<tr>
<td>Error – Pooling</td>
<td>28.35</td>
<td>43</td>
<td>0.659</td>
<td></td>
<td>34.95</td>
<td>7.81</td>
</tr>
<tr>
<td>Total</td>
<td>53</td>
<td></td>
<td></td>
<td></td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

8. OPTIMISATION OF RESULTS – LAG OF ACCOMMODATION

Refer table 8; the response variations for each level factors are,

- **Factors** Response variation
  - A1: 296
  - A2: 273
  - A3: 212
- **Factors** Response variation
  - B1: 286
  - B2: 249
  - B3: 246
- **Factors** Response variation
  - C1: 226
  - C2: 261
  - C3: 294

*Bolded letters indicate the minimum value.

Similarly, the response variations for each level of all possible interactions are also done and we got the same results. On comparing the response variations of all factors and their interactions at all levels, the minimum values are found out. For this case, the optimal combination for minimum eyestrain is A3B3C1, i.e.,

- Viewing angle = -25°
- Viewing Distance = 80 cm
- Exposure Time = 45 min

The same result is also inferred by plotting response variations against the all the levels of factors and their possible interactions indicate the response graphs for individual factors and their interactions.

9. CONFIRMATION EXPERIMENT

A confirmation experiment is performed by conducting a test using a specific combination of the factors and levels previously evaluated. The purpose of the confirmation experiment is to validate the conclusions drawn during the analysis phase. Significant combination: A3B3C1. The table shows the result of the confirmation experiments conducted at the significant level of factors and the all the response variables measured for validation.

<table>
<thead>
<tr>
<th>Trail No.</th>
<th>Lag of accommodation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Before</td>
</tr>
<tr>
<td>1</td>
<td>155</td>
</tr>
<tr>
<td>2</td>
<td>161</td>
</tr>
<tr>
<td>3</td>
<td>158</td>
</tr>
<tr>
<td>4</td>
<td>153</td>
</tr>
<tr>
<td>5</td>
<td>156</td>
</tr>
<tr>
<td>μ\text{obtained}</td>
<td>9.8</td>
</tr>
</tbody>
</table>

10. CONCLUSION

In this work, an attempt has been made to study the relationship between the eyestrain and the factors affect the eyestrain such as viewing angle, viewing distance and exposure time. Also, as attempt has been made to optimise the values of these factors using Taguchi method that will result in minimum eyestrain.

From the results and discussions, we conclude that

- VDT must be viewed at a downward angle of about 25°
- Viewing distance should be 80cm away from the screen and
- There must be sufficient periodic rest breaks after continuous exposure for 45 minutes.

These optimal values have also been validated by the confirmation experiment.

\[
\text{\mu}_{\text{eff}} = \frac{54}{(1+6)} = 7.71
\]

Calculation:

\[
\text{\mu}_{\text{eff}} = \frac{54}{(1+6)} = 7.71
\]

\[
\text{CI} = \sqrt{(F_{2,11} \times \text{var} \left[ \left( \frac{1}{\text{\mu}_{\text{eff}}} \right) \right])}
\]

\[
= \sqrt{7.66 \times 1.0043 \times 0.3296}
\]

\[
= 1.592
\]

\[
\text{\mu}_{\text{predicted}} > \text{\mu}_{\text{obtained}}
\]

International Conference on Recent Trends In Engineering & Technology (ICRTET) 13th Jan 2013, Bangalore, ISBN: 978-93-82208-54-9

11
REFERENCES:


“INTERFERENCE REDUCTION IN MOBILE AD HOC NETWORK USING DIRECTIONAL ANTENNA”

PRADEEP KUMAR PATIL1 & PALLAVI S.BIRADAR2

1Department of Computer Science & Engg, Appa Institute of Engg & Technology, Gulbarga, India
2Department of Information Science & Engg, Poojya Doddappa Appa College Of Engg, Gulbarga, India

Abstract - Adhoc wireless networks are the combination of both mobile nodes and fixed wireless devices. Over the past few years research into adhoc network has yielded considerable advances, notably in the areas of new routing techniques. Yet, the significant shortcoming of adhoc networks remains the same. Hence several routing schemes have been proposed which use multiple paths simultaneously by splitting the information among multitude of paths, as it may help to reduce end-to-end delay and packet loss. With the development of location aware wireless application, location determination has become an increasingly important middleware technology. Numerous current technologies for location determination of a node use the received signal strength from the nodes using Omni directional antennas. However, an increasing number of wireless adhoc systems are now deploying directional antennas due to their advantages like energy conservation, reduced packet loss and better bandwidth utilization. In wireless adhoc network position awareness is necessary to exploit the communication benefits of directional antennas and for nodes to provide meaningful information about their surroundings. The result of the simulation study shows that directional antenna has overcome the traditional Omni directional antenna due to their added advantages as mentioned earlier. With directional antenna on a nodes overall network throughput can be improved, which results into prolonged network lifetime.

Keywords- Directional Antenna.

1. INTRODUCTION

The recent progress in wireless communication and personal computing leads to the research of adhoc wireless networks, which are envisioned as rapidly deployable, infrastructure-less networks with each node acting as a mobile router and equipped with a wireless transceiver. Usually, in adhoc networks all nodes are equipped with Omni directional antenna. However, nodes in adhoc network with Omni directional antenna uses RTS/CTS based floor reservation scheme that wastes a large portion of the network capacity by reserving the wireless media over a large area. To alleviate this problem, researchers have proposed to use directional (fixed or adaptive) antennas that direct the transmitting and receiving beams only towards the receiver and transmitter nodes. This would largely reduce radio interference, thereby improving the utilization of wireless medium and consequently the network throughput [1-4]. In order to fully exploit the capability of directional antenna, it is necessary for each node to know the information of the neighboring nodes (such as node-ID, direction, link quality, power etc) beforehand. It is possible for a node to have up-to-date information of its location if it contains location determination hardware, such as a GPS receiver, mounted on it. There are different reasons why it may not be feasible to equip each node with such hardware. Thus, the combined unit of the adhoc node and the GPS receiver will have to be replaced far too frequently for it to be practicable for a large class of deployments. Location tracking mechanism in the context of wireless adhoc networks with directional antenna is a serious problem, since it incurs a lot of control overhead. In order to track the location of its neighbor, each node n periodically collects its neighborhood information and forms an Angle- Signal Table (AST). Multipath routing permits the establishment of multiple paths between a source and a destination. This can provide various benefits over single path protocols in wireless networks, increasing reliability, reducing power transmission, enabling load balancing and bandwidth aggregation, and secure communications [15]. The knowledge of spatially disjoint routes enables the effective exploitation of Multipath routing. It can helps to avoid radio collisions between alternate paths, avoid regional failures (translates into better reliability) and can provide secure routing. In this paper we make use of an on-demand Multipath protocol based on source routing called Spatial Disjointness in Multipath Routing (SDMR), which is capable of finding spatially disjoint paths.

2. RELATE WORK

In the study of a Mobile adhoc network protocol, it is important to simulate the protocol and evaluate its protocol performance. Protocol simulation has several key parameters, including mobility model and communicating traffic pattern. Mobility models characterize user movement patterns, i.e. the different behaviors of subscribers. Traffic models describe the condition of the mobile services. Guaranteeing Low Bit Error Rate (BER) and intern high packet delivery ratio is one of the most important issues in MANET. These suffer from the
problem of fading and interference in the physical layer which ultimately reduces the power of transmission. The problem can be solved by using directional antenna, which has a main lobe and smaller side lobes. In the main lobe power gain is maximum and thus if a path is selected across the nodes through such a path, it will have minimum loss. But such an ideal path can never be generated. Therefore it is quite obvious that there will be some losses. These losses can be overcome by integrating the loss calculation in routing layer which takes into account of minimum loss (power) path. Therefore the objective of this project is to develop a multipath routing for MANET by considering Directional Antenna with power loss calculation.

3. ROUTING PROTOCOL DESCRIPTION

The main aim is to improve the performance of DSR protocol with directional antenna by taking transmission and received power into consideration. Multiple routes are constructed from source to destination. Among all the obtained routes, the route with the low cost and minimum interference is chosen for transmission. In order to achieve the desired objective, we use network simulator OMNET++ which has a built in framework and which can be modified and work can be implemented. This will save considerable amount of time and effort needed to implement the requirement at hand, providing the desired results. In this work, we propose to reduce the interference in MANET using Directional antenna as an alternative to the existing technique, Omni directional antenna. We estimate the power required for the transmission of data from source to destination in the multipath environment, we consider the path which has low cost and minimum interference.

4. SIMULATION RESULTS AND ANALYSIS

The simulation is carried out for the directional antenna in MANET with the routing layer along with the Omni directional antenna in omnet++ version 3.2. The simulation is carried out by varying nodes. The simulation is carried out for the directional antenna in MANET with the routing layer along with the present system (Omni directional antenna) by varying nodes ranging from 15 nodes to 35 nodes for a period of 1000 sec in express mode and fixing the pause time for 300 sec. We measure the throughput (Mbps), packet delivery ratio (%), control overhead and latency (sec). This simulation is carried out for four sources and four destinations i.e. four sessions. Now we shall compare each and every simulation parameter for both the present and proposed system, this can be achieved by plotting the graph for each parameter.

Graph-1 NODES Vs PACKET DELIVER RATIO

From the graph it is clear that the proposed system, directional antenna has a steady packet delivery ratio when it is compared with the present system, omnidirectional antenna throughout the node density ranging from 15 nodes to 35 nodes. This is because of the fact that the source selects the path which has low cost and less interference.

Graph-2 NODES Vs LATENCY

The above graphs shows that the directional antenna outperforms the omnidirectional antenna in terms of latency. This behavior is more noticeable when the number of nodes increases.

Graph-3 LOAD Vs PACKET DELIVERY RATIO

The graph shows the packet delivery ratio (PDR) of the two different modules such as Omni and proposed Directional antenna. We can observe that PDR of directional antenna is better than omni directional especially when increasing the network size.
Graph-4 LOAD Vs LATENCY

The above graph shows the result of Directional and Omni directional antenna for 4 session, 1000 seconds of simulation, pause time 300 seconds.

Graph-5 PAUSE TIME Vs PACKET DELIVERY RATIO

The graph shows the packet delivery ratio (PDR) of the two different modules such as Omni and proposed Directional antenna. We can observe that PDR of directional antenna is better than omni directional even though there is an increase in pause time.

Graph-6 PAUSE TIME Vs LATENCY

The above graphs shows that the directional antenna out performs the omni directional antenna in terms of latency. This behavior is more noticeable when the number of nodes increases.

5. CONCLUSION

By using directional antenna, a node may be able to selectively receive signals only from a certain desired direction. This enables the receiver node to avoid interference that comes from unwanted direction, thereby increasing the signal to interference and noise ratio (SINR). Directional antenna has a main lobe and small side lobes. In the main lobe, power gain is maximum and the path selected across the nodes will have minimum loss. Multiple paths are generated from Source to Destination at the initial phase of the communication. During the data transmission phase, Source node finds the path with least cost and transmits through that path. The result shows that power loss estimation with Omni-direction antenna is less accurate than with directional antenna, the life time of the network is also increased and latency is very minimal as compared with the Omni directional antennas.

6. ACKNOWLEDGEMENTS

Authors would like to thank Prof. Sujatha Terdal, Department of Computer Science and Engineering, PDA college of Engg, Karnataka for her guidance, constant encouragement and providing the necessary help.

REFERENCE


[10] Location-Aware Power-Efficient Directional MAC in Adhoc Networks using Directional Antenna
“Interference reduction in mobile ad hoc network using directional antenna”


[16] Multipath Routing in Adhoc Networks Using Directional Antennas Yang Li, Hong Man, Jin Yu, Yu-Dong Yao Dept. of Electrical and Computer Engineering Stevens Institute of Technology, Hoboken, NJ 07030

[17] Deafness and Virtual Carrier Sensing with Directional Antennas in 802.11 Networks Michael Neufeld and Dirk Grunwald University of Colorado Department of Computer Science Boulder, CO 80309

[18] Interference in Wireless Multi-hop Ad-hoc Networks Using Omnidirectional and Directional Antennas Bijoy Rahman Arif School of Engineering and Applied Sciences North South University Dhaka, Bangladesh

[19] Adhoc Networking With Directional Antennas: A Complete System Solution Ram Ramanathan, Senior Member, IEEE, Jason Rudi, Senior Member, IEEE, Cesar Santivanez, Member, IEEE, David Wiggins, and Stephen Polit


[21] A Distributed Asynchronous Directional-to-Directional MAC Protocol for Wireless Adhoc Networks Emad Shihab, Student Member, IEEE, Lin Cai, Member, IEEE, and Jianping Pan, Senior Member, IEEE


[23] A Directional MAC protocol for Practical Smart Antennas: Yuya Takatsuka, Katsushiro Nagashima, Masanori Takata, Masaki Bandai, Takashi Watanabe Graduate Schol of Informatics, Shizuoka University.


“FLOW BASED DIRECTIONAL ROUTING WITH LOAD BALANCING FOR MOBILE AD HOC NETWORKS”

PALLAVI S.BIRADAR¹ & PRADEEP KUMAR PATIL²

¹Department of Information Science & Engg, Poojya Doddappa appa College Of Engg, Gulbarga, India
²Department of Computer Science & Engg, Appa Institute of Engg & Technology, Gulbarga, India

Abstract-A mobile ad hoc network (MANET) is an infrastructure less distributed network without any central control, in which each node jointly participates for the routing. Routing in such an infrastructure less network is still a challenging task. Multipath routing protocols are distinguished from single-path routing by the fact that they look for and use several paths to distribute the traffic from source to a destination instead of routing all the traffic along a single path. Multiple paths improve network performance as compared to a single path in the communication. The communicating nodes in ad-hoc network are subjected to co-channel interference. The performance of ad-hoc network is degraded with interference in the network. Thus mobile devices are equipped with Antennas. There are several routing schemes with Directional antenna have been proposed, which uses multiple paths simultaneously by splitting the information among multitude of paths, as it may help to reduce end-to-end delay and perform load balancing. Two main aspects are considered in this paper. The First aspect is to perform preemptive route rediscoveries before the occurrence or route errors while transmitting a large volume of data from Source to Destination. The Second aspect is to calculate the cost of the path considering correlation factor and flow. Among the available paths the first two best routes are selected and the data is split across the selected two paths. The result of the simulation study shows that Flow based directional antenna improves the performance of multipath routing significantly as compared to that with Flow based Omni-directional antenna.

Keywords: Load balancing, correlation factor, flow.

1. INTRODUCTION

An Ad Hoc wireless network is a collection of two or more devices equipped with wireless antenna and networking capability. Such devices can communicate with another node that is immediately within their radio range or one that is outside their radio range with the help of intermediate neighboring node(s). An intermediate node acts as a messenger to relay or forward the packet from the source towards the destination(s). The major part in the area of a Mobile Ad Hoc Network (MANET) has been dedicated to the fundamental issues of topology control and routing. Routing algorithms try to actually route information along optimal paths that are theoretically present due to topology control. Usually wireless networks go for sparse network. The main reason for constructing sparse networks is to reduce interference. Till today no topology control algorithm has been known that explicitly minimizes interference! The objective is to thoroughly define and analyze the notion of interference in ad-hoc networks. To reduce interference among the Mobile devices can be achieved by equipping Antennas.

Directional antennas add new degrees to freedom and operational choices to MAC and network-layer protocols. The concept of using directional antennas for performance enhancements in ad hoc networks is not new. Several routing schemes and MAC layer protocols involving directional antennas have been described. Multipath routing protocols with directional antennas are distinguished from single-path routing by the fact that they look for and use several paths to destination instead of routing all the traffic along a single path.

In order to show the performance of multipath routing with directional antenna through proper load balancing improves substantially, the effect of route coupling is considered which is measured using a correlation factor. In a network there might be many nodes actively sending the data at a time. The flow represents the data packets of the unique source destination pair. By selecting a path with minimum correlation factor and maximum no of flow we can reduce the control overhead and thus control the congestion in the network.

The splitting of traffic in mobile ad hoc network on the different multipath can balance the load in the network. The splitting of traffic can be done equally or based on parameters like delay, energy or bandwidth available. In this paper we demonstrate a unique solution for MANET by introducing the flow based multipath directional routing concept whereby a node participates in the routing if only it satisfies the correlation and flow criteria. Further we obtain multiple paths based on the number of flows served by the nodes. Once the paths are obtained between source and destination pair, we split the traffic according to a cost determined through the sum of correlation factor and number of flows.
2. RELATED WORK

An ad-hoc network is a collection of communicating devices that intends to communicate, but have no fixed infrastructure available, and have no pre-determined organization of available links. It is stated [1] that in a wireless ad-hoc network communication between nodes takes place over radio channels and for this reason it becomes subject to co-channel interference or simply interference. To reduce interference in wireless ad hoc networks, mobile devices are equipped with antennas as proposed in [2] Omni directional antennas transmit and receive electromagnetic signals equally in all directions. It has revealed that ad hoc network with directional antennas increase spatial reuse. A detailed comparative study is made in [3] which state that Directional antennas partition the Omni directional transmission area into a fixed number of sectors. Transmission in one sector will not affect signal propagation in other sectors. Therefore a spatial region previously occupied by one Omni directional transmission may now be shared by several directional transmissions.

Routing schemes are based on the shortest path metric. Directional routing is been considered and its various issues have been discussed [1-5]. Directional routing suffers due to frequent node movements this give rise to problems such as locating and tracking during random channel access[2]. This issue is discussed in [4], where it is assumed that every node knows its direction to other nodes in the network. In [5], it reveals that the directional antennas are used to improve routing performance. However various protocols have been focused on making changes due to the technical and technological challenges involved in already existing directional protocols.

Recent studies extensively focus on multipath extension of the on-demand routing protocols in order to alleviate single-path problems, such as high route discovery latency, frequent route discovery attempts and possible improvement of data transfer throughput. Multipath routing scheme employs a set of paths from source S to destination D so that total volume of traffic may be divided and communicated via selected multiple paths which would perform load balancing and eventually reduces congestion and end to end delay. These various directional routing protocols use multiple paths simultaneously by splitting the information among multitude of paths [7] in a network there might be many nodes actively sending the data at a time. The flow represents the data packets of the unique source destination pair. Any node can be a part of many flow or transmission at a time. A path with minimum no of flow is been selected [8]. Since even load is considered as a routing parameter, a Load –Sensitive routing (LSR)

protocols for static wireless ad hoc networks using directional antennas is presented [6].The optimal solution is obtained by applying the flow based directional algorithm in distributive manner.

3. ROUTING PROTOCOL DESCRIPTION

Dynamic Source Routing (DSR) protocol is used for finding multiple loop-free and link disjoint paths. These routing schemes use multiple paths simultaneously by splitting the information among multitude of paths, as it may help to reduce end-to-end delay and perform load balancing. Multipath routing also diminishes the effects of unreliable wireless links in the constantly changing topology of ad hoc networks to a large extent. Multipath routing scheme employs a set of paths from source S to destination D so that total volume of traffic may be divided and communicated via selected multiple paths which would perform load balancing and eventually reduces congestion and end to end delay. A load-balancing scheme has a significant effect on the performance of the multipath routing protocol, especially in an ad hoc network environment. This DSR protocol is modified for static wireless ad hoc networks using directional antennas. Data delivery - ratio, end –to-end delay, and throughput are used as the performance metric to compare the routing metrics. This modified protocol considers the no of active nodes as the load metric for choosing routes. When any node has data packets to send the other node it first checks whether the routes to the destination in the route cache are present or not. If routes are already present it sends the packets on those paths. If routes are not present then it floods the RREQ message to all its neighboring nodes. When the RREQ message is received by any intermediate node then first it checks if route to destination is known if so forward packet else generates RREQ packet. Using find node we can search a node in the routing table. The CF and Flow value is appended along with the RREQ packet, when RREQ reaches the destination the incoming paths are put in the cache. RREP is generated and if node is part of RREP, it must update its flow value. The first two best routes are selected. Then equally split the traffic in two parts along these selected paths and generate the data message.

4. SIMULATION RESULTS AND ANALYSIS

The simulation is carried out for the directional antenna in MANET with the routing layer along with the Omni directional antenna in omnett++ version 3.2.

- The simulation is carried out by varying nodes ranging from 15 nodes to 35 nodes for a period of 1000 sec in express mode and fixing the pause time for 300 sec. We measure the throughput (Mbps), packet
delivery ratio (%), control overhead and latency (sec).

- Next we vary load ranging from 100 to 350, fixing nodes as 25 for a period of 1000 sec in express mode and fixing the pause time for 300 sec. We measure the throughput (Mbps), packet delivery ratio (%), control overhead and latency (sec).
- Finally, we vary pause time ranging from 150 to 350 sec, fixing nodes as 25 for a period of 1000 sec in express mode and fixing the load for 100. We measure the throughput (Mbps), packet delivery ratio (%), control overhead and latency (sec).

Below graphs shows the simulation carried out by varying nodes.

**NODES Vs PACKET DELIVERY RATIO:**

![Graph 1](image1)

**Graph 1**

From the graph 1 it is clear that the proposed system, directional antenna has a steady packet delivery ratio when it is compared with the present system, Omni-directional antenna throughout the node density ranging from 15 nodes to 35 nodes. This is because of the fact that the source selects the path which has low cost and less interference.

**NODES Vs THROUGHPUT:**

![Graph 2](image2)

**Graph 2**

When we consider the node density ranging from 15 nodes to 35 nodes for understanding the throughput, the proposed system has a better throughput compared with the Omni directional antenna can be viewed from in graph 2.

**NODES Vs LATENCY**

![Graph 3](image3)

**Graph 3**

Graph 3 shows that the directional antenna out performs the Omni directional antenna in terms of latency. This behavior is more noticeable when the number of nodes increases.

**NODES Vs CONTROL OVERHEAD**

![Graph 4](image4)

**Graph 4**

From the above graph 4 we can say that the control overhead of the proposed technique is much lesser than the present technique.

Similarly series of simulation is carried out varying Load from 100 to 350 and Pause Time from 150 to 350 fixing nodes as 25 for a period of 1000 sec in express mode. We measure the throughput (Mbps), packet delivery ratio (%), control overhead and latency (sec).

5. **CONCLUSION**

In this work we introduced a Flow based directional routing algorithm with reduced control overhead for mobile Ad Hoc Networks. In this work we combined flow based directional multipath routing with load balancing mechanism. In flow based directional multipath routing uses the information of number flows served by a node and hence reduce the propagation of control packet while ensuring better distribution of data packets between the nodes in the network. The restricted propagation of RREQ or control packets while route discovery helps in reducing the control overhead and thus reduce occurrence of congestion in network. The splitting of the traffic or load onto multipath based on delay helps to efficiently balance the load and hence ensures the proper utilization of the network resources. From the simulation run above obtained graphs we can conclude that the throughput and PDR are significantly improved in the proposed work when
compared to the present Omni directional multipath routing protocols. The control overhead is reduced efficiently in our system when compared to the other multipath routing protocols. The proposed system works well for both smaller and larger size networks with many numbers of nodes with many active sources.

6. ACKNOWLEDGEMENTS

Authors would like to thank Prof. Sujatha Terdal, Department of Computer Science and Engineering, PDA college of Engg, Karnataka for her guidance, constant encouragement and providing the necessary help.

REFERENCES


[6] Load Balancing of Multipath Source Routing in Ad Hoc Networks. Lianfang Zhang, Zenghua Zhao, Yantai Shu, and Lei Wabg Department of computer science Tianjin University, Tianjin 300072, China. Oliver W.W.Yang, School of information Technology and Engineering University of Ottawa


PERFORMANCE ANALYSIS OF VARIOUS EDGE DETECTION TECHNIQUES ON LUMBAR SPINE IMAGE

E.PUNARSELVAM¹ & P.SURESH²

¹Department of Information Technology, Muthayammal Engineering College M, Rasipuram – 637 048, Tamilnadu, India.  
²Department of Mechanical Engineering, Karpagam College of Engineering, Coimbatore– 641 032, Tamilnadu, India.

Abstract - The lumbar vertebrae are the largest segments of the movable part of the vertebral column, they are elected L1 to L5, starting at the top. The spinal column, more commonly called the backbone, is made up primarily of vertebrae discs, and the spinal cord. Acting as a communication conduit for the brain, signals are transmitted and received through the spinal cord. It is otherwise known as vertebral column consists of 24 separate bony vertebrae together with 5 fused vertebrae, it is the unique interaction between the solid and fluid components that provides the disc strength and flexibility required to bear loading of the lumbar spine. In this work the performance comparison between lumbar spine edge detection algorithms based on Magnitude and Edge Length of CT scan spine disc image. The result shows that the canny edge detection algorithm produced better result when compared OTHER edge detection algorithm. Finding the correct boundary in a noisy image of spine disc is still a difficult one. To find out absolute edges from noisy images, the comparative result can be verified and validated with the standard medical values. The result shows that the canny edge detection algorithm performs well and produced a solution very nearer to the optimal solution. This method is vigorous for all kinds of noisy images.

Keywords- Spine Disc image, Finite Element Modelling, Sobel Edge Detection, Prewitt operator, Canny Edge Detection, CT scan, Magnitude and Edge Length

I. INTRODUCTION

The key functions of the Human spine as a composite structure is to protect the spinal cord. Manual material handling operations are carried out in most industrial plants. Each behaviour task poses unique trouble on the employee. However, workplaces can help employees to perform these tasks safely and easily by implementing and maintenance proper policies and procedures. The lumbar region which is one of the parts in the spine has played a vital role in the researches. The anatomy lumbar region, the lumbar spine, the back pain and their relationship are put together in literature. The further analysis of the spine under the aircraft ejection is made and this deals with the FEM modelling of the spine for aircraft injuries [1, 2]. The finite element analysis in the field of orthopaedics for the lumbar region [3] has also been explained with its uses. The evaluation and management of occupational low back disorders and back pains were studied. The movement of human by motor control and biomechanics [4, 5] are illustrated in various aspects. Such similar investigations are made on the spine disk and are reported. Till date, the motion segment of the spine consists of two parts of which one is a vertebrae and the other one a disc [6,7,8]. To explicate proper information from the images, a fundamental tool used is the Edge detection technology. This provides the outlines of an image and boundaries [9]. This also proves as a tool to remove the noises in the images to enhance the appearance. The magnitude and Edge length based algorithm called the Canny Edge detection algorithm has been proposed for pre-processing the boundary detection of the CT-scanned disk image of the spine. To detect the correct boundary in a noisy image is still difficult. Based on the evaluation the canny edge detection algorithm is used to detect the boundaries of spine disc image from the noisy CT scanned image produce a better result. The Fig.1 shows a 2D model of spine disc image and the Fig.2 gives a clear view of the same in 3D model. The vertebral body of each lumbar vertebra is large, wider from side to side than from front to back, and a little thicker in front than in back. It is flattened or slightly concave above and below, concave behind, and deeply constricted in front and at the sides. The Fig.3 illustrates the various spine disc disorders. And Fig. 4 represents the spine image, and
Table 1 gives the behaviours and properties of the spine. FEA consists of a computer model of a material or design that is stressed and analysed for exact results. It is used in new manufactured goods design, and existing product refinement. A company is able to verify a proposed design will be able to perform to the client’s specifications prior to manufacturing or construction. Modifying an existing product or structure is utilized to qualify the product or structure for a new service condition. In case of structural failure, FEA may be used to help determine the design modifications to meet the new condition.

II. VARIOUS EDGE DETECTION TECHNIQUES

APPLIED FOR LUMBAR SPINE DISC IMAGE

- Three most frequently used edge detection methods are used for comparison that is (1). Sobel Edge Detection, (2). Prewitt Edge Detection and (3). Canny Edge detection. The detail of each method as follows,

2.1 PROPOSED WORKFLOW SEQUENCE
2.2. Sobel Edge Detection: The Sobel operator performs a 2-D spatial gradient measurement on an image and so emphasizes regions of high spatial frequency that correspond to edges. Typically it is used to find the predictable absolute gradient magnitude at each point in an input gray scale image. The operator consists of a pair of 3x3 convolution kernels as shown in Figure 6. One kernel is simply the other rotated by 90°. This is very similar to the Roberts Cross operator[12]. The convolution masks of the Sobel detector are given

\[
G_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} \quad G_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}
\]

These kernels are designed to respond maximally to edges running vertically and horizontally relative to the pixel grid, one kernel for each of the two perpendicular orientations. The kernels can be applied separately to the image, to produce separate measurements of the gradient component in each orientation (\(G_x\) and \(G_y\)). These can then be combined together to find the absolute magnitude of the gradient at each point and the orientation of that gradient. The gradient magnitude is given by

\[
|G| = \sqrt{G_x^2 + G_y^2}
\]

Typically, an approximate magnitude is computed using:

\[
|G| = |G_x| + |G_y|
\]

which is much faster to compute. The angle of orientation of the edge (relative to the pixel grid) giving rise to the spatial gradient is given by:

\[
\theta = \arctan(G_y/G_x)
\]

in this case, orientation 0 is taken to mean that the direction of maximum contrast from black to white runs from left to right on the image, and other angles are measured anti-clockwise from this. Often, this absolute magnitude is the only output the user to the two components of the gradient are conveniently computed and added in a single pass over the input image using the pseudo-convolution operator shown in Figure 7.

\[
\begin{array}{ccc}
P_1 & P_e & P_o \\
P_s & P_5 & P_p \\
P_p & P_s & P_p
\end{array}
\]

2.3. Prewitt Edge Detection: The prewitt edge detector is an accurate way to approximation the magnitude and orientation of an edge. Although differential gradient edge detection needs a rather time consuming computation to estimate the orientation from the magnitudes in the x- and y-directions, the compass edge detection obtains the orientation directly from the kernel with the maximum response [13,14]. The prewitt operator is limited to 8 possible orientations, however experience shows that most direct orientation estimates are not much more accurate. This gradient based edge detector is estimated in the 3x3 neighbourhood for eight directions. All the eight convolution masks are calculated. One convolution mask is then selected, namely that with the largest module[17]. Mathematically, the operator uses two 3x3 kernels which are convolved with the original image to calculate approximations of the derivatives - one for horizontal changes, and one for vertical. If we define \(A\) as the source image, and \(G_x\) and \(G_y\) are two images which at each point contain the horizontal and vertical derivative approximations

\[
G_x = \begin{bmatrix} -1 & 0 & +1 \\ -1 & 0 & +1 \\ -1 & 0 & +1 \end{bmatrix} \quad \text{and} \quad G_y = \begin{bmatrix} +1 & +1 & -1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}
\]

where \(*\) denotes the 2-dimensional convolution operation. Since the Prewitt kernels can be decomposed as the products of an averaging and a differentiation kernel, they compute the gradient with smoothing. \(G_x\) can be written as

\[
\begin{bmatrix} -1 & 0 & +1 \\ -1 & 0 & +1 \\ -1 & 0 & +1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}
\]

Results At each point in the image, the resulting gradient approxim approximations can be combined to give the gradient

\[
G = \sqrt{G_x^2 - G_y^2}
\]

Using this information, we can also calculate the gradient's direction:

\[
\Theta = \arctan(G_y/G_x)
\]

Where \(\Theta\) is 0 for a vertical edge which is darker on the right side. Prewitt operator at an image point which is in a region of constant image intensity is a zero vector and at a point on an edge is a vector which points across the edge, from darker to brighter values.

2.4. The Canny edge Detection:

The Canny algorithm can be used an optimal edge detector based on a set of criteria which include finding the most edges by minimizing the error rate, marking edges as closely as possible to the actual edges to maximize localization, and marking edges...
[10,11] only once when a single edge exists for minimal response. According to Canny,

Table 2. Comparison Between Various Edge Detection Techniques for CT Scan Lumbar Spine Disc Image.

<table>
<thead>
<tr>
<th>No</th>
<th>Image Method</th>
<th>Medical Value (%)</th>
<th>Sobel Edge Segmentation (%)</th>
<th>Canny Edge Segmentation (%)</th>
<th>Error Difference in Sobel Edge Segmentation (%)</th>
<th>Error Difference in Canny Edge Segmentation (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CT Scan Disc image</td>
<td>8.51</td>
<td>8.55</td>
<td>8.60</td>
<td>8.52</td>
<td>-0.04</td>
</tr>
<tr>
<td>2</td>
<td>Average magnitude in up</td>
<td>7.73</td>
<td>7.79</td>
<td>7.79</td>
<td>7.76</td>
<td>-0.06</td>
</tr>
<tr>
<td>3</td>
<td>Density of edge length in up</td>
<td>7.90</td>
<td>7.97</td>
<td>7.94</td>
<td>7.91</td>
<td>-0.07</td>
</tr>
<tr>
<td>4</td>
<td>Blurred CT Scan image</td>
<td>6.44</td>
<td>6.26</td>
<td>6.26</td>
<td>6.42</td>
<td>-0.09</td>
</tr>
<tr>
<td>5</td>
<td>Initial position map</td>
<td>5.99</td>
<td>5.92</td>
<td>5.98</td>
<td>5.92</td>
<td>-0.06</td>
</tr>
<tr>
<td>6</td>
<td>Pre Edge filtered in up</td>
<td>5.17</td>
<td>5.14</td>
<td>5.19</td>
<td>5.12</td>
<td>-0.07</td>
</tr>
</tbody>
</table>

The optimal filter that meets all three criteria above can be efficiently approximated using the first derivative of a Gaussian function.

\[
G(i,j) = \frac{1}{2\pi\sigma^2} e^{-\frac{(i^2 + j^2)}{2\sigma^2}}
\]

**III. PRE-PROCESSING OF INITIAL POSITION OF EDGE PARAMETERS DETECTION**

**Step 1:** Calculate the average magnitude

\[
M(1,2) = \frac{1}{n} \sum_{i=1}^{n} \sqrt{L(1,2)^2 + M(1,2)^2}
\]

**Step 2:** Calculate the density of the edge length

The density of the edge length is calculated from

\[
L(1,2) = \frac{C(i,j)}{\text{max}(C(1,2))}
\]

Where \(C(i,j)\) is the number of connected pixels at each position of pixel.

**Step 3:** Calculate the initial position of map from summation of density of edge length and average magnitude.

\[
P(1,2) = \frac{1}{\sqrt{M(1,2) + L(1,2)^2}}
\]

**Step 4:** Calculate the thresholding of the initial position map. If \(P(1,2) > T_{\text{max}}\)
Then \(P(1,2)\) is the initial position of the edge following. And then we obtained the initial position by setting \(T_{\text{max}}\) to 92% of the maximum value.

**IV. RESULTS AND DISCUSSION**

To further evaluate the efficiency of the proposed method in addition to the visual inspection, the proposed boundary detection method numerically using the Hausdorff distance and the probability of error in image segmentation. Where \(P(O)\) and \(P(B)\) are probabilities of objects and background in images. The objects surrounded by the contours obtained using the five snake models and the proposed method are compared with that manually drawn by skilled doctors from the Medical Hospital. From the above Table.2 shows the average result of probability of Error in Image segmentation of sobel, Prewitt operator and canny edge detection algorithm were compared with standard Medical values and also predicts the error difference. Showing the results it shows the Error difference value is very minimal and also negligible in Sobel edge detection algorithm. So the Canny edge detection algorithm produced nearer to the standard value. Fig.8 Shows the comparative analysis of Sobel edge detection and Medical standard value. Fig.9 Shows the comparative analysis of Prewitt operator edge detection and Medical standard value. Fig.10 Shows the comparative analysis of Canny edge detection and the Medical standard value which is collected from the standard Hospital.

![Fig 8. Comparative Analysis Graph for Sobel Edge Detection value and Medical Standard value.](image)
In this paper, the comparative analysis of various edge detection algorithm like sobel, Prewitt and canny edge detection algorithm out of which canny edge detection algorithm produced better result for noisy spine disc image. So based on the result the proposed technique for boundary detection is applied it to spine disc image. Our edge following technique incorporates a vector image model and the edge map information. The proposed technique was applied to detect the object boundaries of noisy CT scan spine disc image where the well defined edges were encountered. The opinions of the skilled doctors were used as the ground truths of interesting object of spine disc. Besides the visual inspection, the canny edge detection method was verified and evaluated using the probability of error. The results of detecting the object boundaries in noisy images show that the canny technique is very close to the standard value which was given by eminent doctors. We have successfully applied the edge following technique to detect the object boundaries of spine disc image. The proposed method can be applied not only for medical imaging, but can also be applied to any image processing problems.

V. CONCLUSION

In this paper, the comparative analysis of various edge detection algorithm like sobel, Prewitt and canny edge detection algorithm out of which canny edge detection algorithm produced better result for noisy spine disc image. So based on the result the proposed technique for boundary detection is applied it to spine disc image. Our edge following technique incorporates a vector image model and the edge map information. The proposed technique was applied to detect the object boundaries of noisy CT scan spine disc image where the well defined edges were encountered. The opinions of the skilled doctors were used as the ground truths of interesting object of spine disc. Besides the visual inspection, the canny edge detection method was verified and evaluated using the probability of error. The results of detecting the object boundaries in noisy images show that the canny technique is very close to the standard value which was given by eminent doctors. We have successfully applied the edge following technique to detect the object boundaries of spine disc image. The proposed method can be applied not only for medical imaging, but can also be applied to any image processing problems.

REFERENCES


ON CATEGORIES OF FACTORS INFLUENCING THE INNOVATION: 
AN EMPIRICAL STUDY IN IT SECTOR

M VISHWANATH PAI¹ & LEWLYN L.R. RODRIGUES²

¹Department of Computer Science, MGM College, Udupi, 576 102, INDIA.
²Humanities and Social Sciences, Manipal Institute of Technology, Manipal, 576 104, INDIA

Innovation has been one of the most important factors in driving the economic growth in all industries and in particular the IT industry. The innovation will give a competitive edge to a company over the others. But the factors which influence innovation are not understood clearly. Nevertheless there are no proper instruments which measure the performance of innovation owing to the complexity involved in selection of different variables. There are some instruments available for innovation performance measurement but they are not complete. The main thrust of this paper is on the development, validation and application of the categories of factors which influence innovation in the IT sector through empirical means. The instrument that is chosen for the empirical study tests the dependence of innovation in the IT sector on the three dimensions of categories of factors. The instrument is validated for content, criterion and constructs validity, which includes the statistical test of significance. The empirical study has been performed using this instrument through random sampling technique. A hypothesis is tested to study whether different categories of factors influence innovation in IT sector. The result of this empirical study is the development of a valid instrument that can be used to study the influence of categories of factors on innovation in IT sector.

1. INTRODUCTION:

Innovation has long been recognized as an important driver of economic growth (Romer 1990, Grossman and Helpman 1994, Bloom and Van Reenen 2002, Bosworth and Collins 2003). Despite the abundance of empirical findings and the unprecedented interest in innovation, researchers still lack a fundamental understanding of the factors that create innovation and the mechanisms through which innovation creates growth. Perhaps most frustrating has been the failure to find an empirical measure of innovative activity that offers deep insight into the underlying factors and mechanisms. In the flurry of theoretical and empirical investigations, most researchers have used intangible assets and total factor productivity growth as proxies for innovative activities. These studies have consistently shown innovative activities to be a major factor in the economy. For example, Nakamura (2001) estimated investment in intangibles at approximately $1 trillion per year, and Corrado, Hulten, and Sichel (2006) estimated those investments at approximately $1.2 trillion per year. The sheer size of these proxy measures indicates the importance of innovation in driving economic growth, but they do not give enough insight into the underlying mechanisms of how innovation yields growth to advance theories and models of innovation.

So it is very important now to identify and categorize different factors which significantly influence innovation in general and IT sector in particular. The IT sector is highly knowledge intensive and needs to be innovative always. It is also very important to find out whether the different categories of factors significantly influence innovation.

2. LITERATURE REVIEW

Despite the difficulties in measuring innovation, researchers persist in their search for the one true indicator of innovation. Milbergs and Vonortas (n.d.) portrayed innovation metrics as evolving through four generations:

• The first generation of metrics reflected a linear conception of innovation focusing on inputs such as R&D investment, and the like.
• The second generation complemented input indicators by accounting for the intermediate outputs of S&T activities.
• The third generation focused on a richer set of innovation indicators and indexes based on surveys and integration of publicly available data.
• The fourth generation metrics of the knowledge-based networked economy remain ad hoc and are the subject of measurement.

Since intangible assets do not exist in physical form, they present a set of difficult measurement problems that arise primarily from the inability to measure intangible assets directly. Researchers are forced to resort to proxies and techniques for indirect measurement. To guide those indirect techniques, researchers have devised a variety of ways to characterize intangible assets. Intangible assets can be divided into three subcategories based on the degree to which they can be controlled and/or sold by the firm (Blair and Wallman 2001):

• Assets that can be controlled and owned by the firm and can be separated and sold, for example, patents and databases.
• Assets that can be controlled and owned by the firm but not separated out and sold, for example, R&D and organizational processes.
on categories of factors influencing the innovation: An empirical study in it sector

- Assets that may not be wholly controlled by the firm and are therefore not owned by the firm, for example, knowledge and skills of labor force.

- While intangible assets represent the knowledge and skill sets of the organization, they are the vehicle for integrating knowledge into an innovative product, service, or process.

3. THE MEASUREMENT INSTRUMENT

The literature review has identified that the intangible assets and tangible assets have significant influence on innovation. These intangible and tangible assets can be grouped into different categories depending on their inherent characteristics. The objective of this research is to find out whether the different categories of factors which have significant influence over innovation are dependent or not.

The meta-analysis of available literature has resulted in the identification of the following instruments used for measuring innovation capacity and intangible assets.

a) **Intangible Asset Monitor**: The Intangible Asset Monitor is a framework for measuring intangible assets and knowledge flows using non-monetary metrics (Bontis 2001; Sveiby 1997). It is based on the premise that firms accumulate intangible assets to enable knowledge and tangible inputs to be converted into tangible outputs and financial outcomes.

b) **Skandia Navigator**: Skandia, a Swedish financial and accounting services firm, was the first large company to produce an intellectual capital report as an addendum to traditional financial reports. Numerous other companies have adopted Skandia’s methods for measuring and reporting intellectual capital (Bontis, 2001).

c) **IC-dVal**: IC-dVal applies a resource-based view of the firm, in which resources are accumulated and deployed through organizational processes to produce outputs and achieve economic outcomes, to correlate the financial value of intangible assets with economic performance (Bounfour 2003). The IC-dVal focuses on four intangible resources:

d) **Cash Curve**: A cash curve illustrates the cumulative flow of cash throughout the innovation process. It uses cash realized (referred to as payback) as a metric for evaluating the progress and success of the innovation process for a specific product. The cash flow at any point in the innovation process is a function of prior investments, current costs, and real and projected sales revenue (from the product). The cash curve provides estimates of cash flow based on assumptions about technical feasibility and the market.

e) **Technology Factors**: The Technology Factors is a tool for managing intellectual assets and evaluating the productivity of R&D and other activities that generate intellectual capital (Bontis 2001). The use of patents and other forms of intellectual property as proxies for intellectual capital is based on empirical evidence that firms with highly cited patents have higher market value.

**Categories of factors**

Thus it is understood that innovation is influenced by a series of factors. These factors are complex, nonlinear, multidimensional, and unpredictable. No single measure is likely to characterize innovation adequately in its totality. Further, important aspects of innovation such as knowledge cannot be measured directly. These factors may be related to human resource, organization, business environment etc. Since the innovation not solely depends on tangible assets the role of intangible assets are also very important.

An organization may have best of the R & D infrastructure but the utilization of it is needed to get the best result. This R & D infrastructure should not be meant to reinvent the wheel but to invent something new. The utilization of the R & D is dependent on different intangible factors predominantly the human capital. New ideas must be generated from within and outside the organization and it should not be detrimental to the society. When an organization spends on R & D or invest on new ideas it may so happen that the organization may lose the investment but the organization should be in a position to sustain such investment be able to absorb the loses and further support the innovation. In such cases the size of the organization is very important and can still sustain the growth and invest on innovative ideas.

The leadership and policies of the organization is very important in promoting innovation. The vision and the direction provided by the top management are very important for an organization to be innovative. The organization must be able to identify the market needs and market trends so as to capitalize on it with the innovative products and ideas. Innovation should be really a need for the organizations which wants to stay ahead in the competition to have an edge over the others. Some organization needs to innovate to
come out of the stagnation. Profit motive of the organizations also make them to be more innovative. Thus there are different reasons for which the companies wish to be innovative.

The economic activity of the region is very important to promote the innovation in that region. The buying capacity of the customers and acceptance of it also is helpful in promoting innovation.

Thus there are various factors which influence innovation in the IT sectors. These various factors are needed to be identified and these different factors may also be independent or dependent on other factors. Since most of these factors are non tangible in nature measuring it will be a difficult task. The different tools considered above will not specify the different categories of factors which have significant influence on innovation are dependent or not.

So the intangible and tangible assets are categorized into three categories as given below.

Category I- Factors related to personal characteristics of an employee
Category II- Factors related to organizations
Category III- Factors external to the organization and employee

Category 1: Personal characteristics
Education, Experience, Reluctance from the employees, Professional Capability, Attrition rate, Mobility of employees, Ideas, Inventions, Communication, Team spirit, Recognition, Motivation

In this category personal characteristic of an IT professional has been considered. The professional can be an employee or can be in the management cadre or the promoter of an IT company itself. Some of these personal characteristics may be unique to an individual person but some may be common with all the professionals. Some of the characteristics may be independent and some may be dependent on others. Again magnitude of this dependency and independency varies on different situation. These personal characteristics may also vary depending on the hierarchy in which the professional is in. The quantum of these characteristics also varies from person to person. But these characteristics do have great influence on innovation.

Category 2: Organization
Top Management Strategy, Human Resource/Knowledge Management, Profitability, Investments/funds, Customer Satisfaction, Forecasting the growth, Research and Development, Size of the company, Pace of reaction to the market needs, Market leadership, Diversity of designs, Sales, Interaction among the different departments, Organizational structure, Product Life cycle, Incentives.

An organization may have its own strength and its own weakness also. Both of these fundamentals depend on many factors. These factors may be dependent or independent of each other. Inherent strength or the weakness of an organization depends not only on the employee but on various other factors. These factors are listed in this category of characteristics. The magnitude of influence of these characteristics on innovation also varies depending on situation and its significance on a given condition. Thus one cannot specify the magnitude of impact of any specific characteristics on innovation but can definitely generalize the quantum of influence. These characteristics are also influenced very much by the personal characteristics.

Category 3: External factors

There are other characteristics which influence innovation may not be related either to personal characteristics of an IT professional or the characteristics related to the organization. But it can be external to both of the categories. Even if both of the above given categories of characteristics have positive impact the third category may have negative impact on innovation. Like other two categories of characteristics different factors of this third category will have varying degree of influence on innovation. Again the magnitude of influence varies on different scenarios and different conditions. These categories have been developed into self administered questionnaire. This questionnaire is used as the measurement instrument and allows holistic assessment of the category of factors which influence innovation activity of a given organization. Depending on different conditions the influence of different category on innovation will vary. The different categories may have varying degree of influence on innovation. The magnitude of influence varies on different counts. Again the magnitude of influence also depends on the category which is being considered on a given situation. Influence of different categories also varies on different scenario. Some of the factors of a category may have different perceptions in different situations. So the magnitude of influence of a category on innovation depends on
the factors whose influence varies on different situations.

The degree of influence may vary from minimum-moderate-maximum. Since each of the categories has different factors in it, influence of each factor will contribute its share to the total influence of each category. When the influence of the category is considered it is the total influence of different factors is considered and not the sole influence of any factors. Influence of different factors of a category also varies on different conditions. Some of these variables of a category are dependent and some are independent. The dependency of a variable on another variable varies in different scenario. Even the independent variables level of influence on innovation is not fixed.

In some conditions the tangible assets play a major role on innovation and in some other cases the intangible assets. Thus again the innovation is not just dependent on tangible and intangible assets but it is the combination of the both. So the organizations can not just invest more on one kind of assets that is tangible or intangible and there should be balance between the two kinds. The proportion of these two kinds of asset is also not the same in all situations and again depends on various factors. Understanding this complex relationship between the different kinds of tangible and intangible assets and the kind of influence these have on innovations is also very important to know which kind of right situations may lead to innovation. But knowing the right situation is also important to get the correct mix of investment on tangible and intangible assets.

In some situations tangible assets play major role and in some situation intangible assets play a major role in innovation. Again the life cycle of the organization in which it is at present also play an important role innovation. There are many companies which are in its initial stage of life cycle innovate and create some kind of ripples in the market through their innovations and will become the market leaders in their segment in the later stage of their life cycle for example Apple, Microsoft, Google, Face Book etc. These companies will move through different stages of their life cycle from initial stage to become the established companies but all these will happen through continuous innovation. There are numerous companies which could not innovate will die down in the initial stages of their life cycle itself and do not move to the next level of life cycle. The successful companies that have already in their advanced stage of life cycle have to continuously innovate to keep the edge over the others. The role of tangible and intangible asset is not the same during different stages of life cycle of a company. For the companies which are in their initial stages of life cycle intangible assets play a major role over tangible. For example companies like Apple, Microsoft, Google, Face Book etc. are the brain child of their founders where the intangible asset of the founders like knowledge, attitude, motivation, ideas etc. have played a major role in innovations. When the companies are in their later stages of life cycle it is not only the intangible but also the tangible asset play a major role in innovation. The position of the company is able to attract more talent who can be offered different kinds of incentives to get the best out of them and retain them too. Through the acquisition routes also the established companies can buy the potential companies which have the capabilities of producing innovative products. The smaller companies which have been sold also get the much needed support of intangible and tangible assets. Since companies which have bought those companies are in the advanced stage of life cycle they can consolidate their positions by innovations being generated by the newly acquired companies.

4. RESEARCH METHODOLOGY:

The sampling unit comprise IT Sector (Software development and Services organizations). The IT Sector is highly knowledge intensive and needs to be innovative always. The sample selection is through random sampling technique. The data collected is through self administered questionnaires that measure the performance effectiveness rating of the three categories. The questionnaire were served to the employees (different levels) of different companies, out of which 131 responses were received back (return rate of 60%) after repeated follow up. However, we could select only 127 entries, as others were incomplete.

The data is processed through Microsoft Excel 2003 for statistical analysis. The processed data is presented in table 4.1 and figure 4.1, 4.2 and 4.3. The empirical study consists of Chi-Square test for independence of attributes in a contingency table. The following hypothesis is tested in this research.

We formulate the Hypothesis H0: Categories and their influence on Innovation are independent. Vs H1: They are dependent.

Table below gives the influence of different identified categories on Innovation

| Table 4.1 |
|------------------|---------------|------------------|------------------|
| Category 1        | Maximum effect on innovation | Moderate effect on innovation | Minimum effect on innovation |
| Catego 33        | 07             | 09               | 49               |
| Catego 25        | 08             | 09               | 42               |
| Catego 12        | 14             | 10               | 36               |
on categories of factors influencing the innovation: An empirical study in it sector

<table>
<thead>
<tr>
<th>Category</th>
<th>Influences related to personal characteristics on innovation</th>
<th>Influences related to organization on innovation</th>
<th>Influences of external factors of business to the organization and employee on innovation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum</td>
<td>33</td>
<td>25</td>
<td>12</td>
</tr>
<tr>
<td>Moderate</td>
<td>7</td>
<td>8</td>
<td>14</td>
</tr>
<tr>
<td>Minimum</td>
<td>9</td>
<td>9</td>
<td>10</td>
</tr>
</tbody>
</table>

5. RELIABILITY, VALIDITY AND PRACTICALITY OF THE MEASUREMENT INSTRUMENT

Sound measurement must meet the tests of Reliability, Validity and Practicality. These are the three major considerations used in a research, which involves data collection through instruments such as questionnaires (Kothari, 2000).

The chi-Square test is an important test amongst the several tests of significance developed by statisticians. Chi-Square test symbolically written as $\chi^2$ test (pronounced as Ki-Square), is a statistical measure used as a nonparametric test, it “can be used to determine if categorical data shows dependency or the two classifications are independent. It can also be used to make comparison between theoretical populations and actual data when categories are used. Thus chi-Square test is applicable in large number of problems. The test is, in fact, a technique through the use of which it is possible for all researchers to (i) test the goodness of fit (ii) test the significance of association between two attributes and (iii) test the homogeneity or the significance of population variance. A common approach used in literature to study the independence of attributes is through Chi-Square test of independence of attributes in a contingency table. The categorized data is represented in a 3X3 contingency table (table no.4.1) and then the above test is applied on it. Here we use this test to test whether the Categories and the influence on innovation are independent (Kothari, 2011).

Practicality of the measurement instrument was evaluated in terms of economy, convenience and interpretability. More items in a questionnaire will give greater reliability (Kothari, 2000) but this is tedious and time consuming. Because of this a limited numbers of items with a maximum of three categories were used in the questionnaire of this research.

6. DATA ANALYSIS AND FINDINGS

The table value of $\chi^2$ for 4 degree of freedom at 5 percent level of significance is 9.488, and the calculated value of $\chi^2$ is 11.31694

Since the calculated value is in the Critical region we reject $H_0$. In other words it may be concluded that the effect on Innovation depends on the categories.

7. CONCLUSION

Innovation in the IT sector has become an important factor for driving this knowledge intensive sector into new heights. But unless these organizations develop an effective measurement instrument to evaluate the performance of innovation, to find whether different categories of factors significantly influence innovation there would be a possibility of loss in productivity and performance and it may adversely affect the return on investment. The literature review has identified that the intangible and tangible assets have significant influence on innovation and has also revealed that measuring of influence of tangible and intangible assets on innovation in terms of revenue generation is difficult and it is not the same always for same set of tangible and intangible assets and again it differs depending on different conditions. Identification of different tangible and intangible asset is an important task. Once these have been identified it is not necessary that they are applicable in the same way in all situations. Some of the factors...
which make up tangible assets may be independent or dependent on other factors. Similarly there may be dependency between the some of the factors which are part of intangible assets and some may be altogether independent. These dependency and independency may again vary on different situations. The quantum of dependency is also not the same all the time and again it may vary during different scenario.

Magnitude of influence of tangible and intangibles assets also vary during the different stages of establishment of an organization. In some stages some of the factors which make intangible assets play a major role and in some cases it is the factors of intangible assets. When there are different factors which influence innovation segregating these factors into different group also a challenging task. Once these factors are grouped these have to be properly categorized into different categories. The factors identified are categorized according to different properties of these factors. The properties of the factors of a group should share some common values.

The statistics have shown that different levels of employees have varying perception about the categories which have significant influence on innovation. According to this perception all these categories have various levels of influence on innovation. In each of the categories the respondents have also stated their perception about the quantum of influence of each of the category on innovation. Thus the respondents have indicated that the influence of each category varies from minimum to moderate to maximum on innovation. Thus again it is the scenario which impress upon the respondents to judge the quantum of influence of a category on innovation. These different scenarios can be related to the tangible or intangible assets. The quantum of influence of tangible and intangible assets may also vary depending on the stage of life cycle in which the company is located at present. Since each of these categories is composed of different factors, each of the factors contribute to varying degree of influence on innovation and the perception of the respondents about the influence of each of the factors on innovation contributes to their view about the kind of influence a category has on innovation.

This empirical study was carried out to provide a valid instrument to find whether different categories of factors significantly influence innovation. A further study on how the different characteristics of intangible and tangible assets influencing innovative on IT sector is in progress. This paper has developed, validated and tested an innovation measurement tool which tests the different categories of factors which significantly influence the innovation in IT sector. There is a scope to further develop this questionnaire to include more number of items to suit the specific requirement of a typical organizational set-up. This tool can also be applied to any kind of industry as the basic components of innovation remain the same. This paper has resulted in the development of one such measurement tool which will readily find the relationship between the different categories of factors which significantly influence innovation.

REFERENCES

DATA WAREHOUSING APPLICATIONS: AN ANALYTICAL TOOL FOR DECISION SUPPORT SYSTEM

MOHAMMED SHAFEEQ AHMED

Department of Computer Science, Gulbarga University, Gulbarga, Karnataka, India

Abstract—Data-driven decision support systems, such as data warehouses can serve the requirement of extraction of information from more than one subject area. Data warehouses standardize the data across the organization so as to have a single view of information. Data warehouses (DW) can provide the information required by the decision makers. The data warehouse supports an on-line analytical processing (OLAP), the functional and performance requirements of which are quite different from those of the on-line transaction processing (OLTP) applications traditionally supported by the operational databases. Data warehouses provide on-line analytical processing (OLAP) tools for the interactive analysis of multidimensional data of varied granularities, which facilitates effective data mining. Data warehousing and OLAP have emerged as leading technologies that facilitate data storage, organization and then, significant retrieval. Both are essential elements of decision support, which has increasingly become a focus of the database industry.

This paper provides a detailed picture of Data warehousing (DW), exploring the features of it, applications and the architecture of DW over Data Mining, Online Analytical Processing (OLAP), On-line Transaction Processing (OLTP) technologies.

Keywords: Data Warehousing, Data Mining, OLAP, OLTP, Decision Making and Decision Support & Managerial decision making.

I. INTRODUCTION

Data Warehousing is a collection of decision support technologies, aimed at enabling the knowledge worker (executive, manager, analyst) to make better and faster decisions for smooth functioning of the enterprise. It serves as a physical implementation of a decision support data model and stores the information on which an enterprise needs to make strategic decisions. It provides architecture and tools for business executives to systematically organize, understand and use their data to make strategic decisions.

Data Warehouse is a database of historical, summarized and consolidated data, which is more important than detailed, individual records used for reporting and analysis. It refers to the database that is maintained separately from an organization’s operational databases which contains consolidated data. The data stored in the data warehouse is uploaded from the operational systems, over potentially long time they tend to be much larger than operational databases. DW systems allow for the integration of a variety of application systems. They support information processing by providing a solid platform of consolidated historical data for analysis.

Bill Inmon (father), the world-renowned expert, said that the definition for a Data Warehousing was and still is today. “A source of data that is subject-oriented, integrated, nonvolatile, and time-variant for the purpose of management’s decision processes”.

Sean Kelly, defines the data warehouse in the following way. “The data in the data warehouse is: Separate, Available, Integrated, Time stamped, Subject oriented, Nonvolatile, Accessible”.

Most queries on data warehouses are ad hoc and are complex queries that can access millions of records and perform a lot of scans, joins, and aggregates. Due to the complexity query throughput and response times are more important than transaction throughput.

A DW (or smaller-scale data mart) is a specially prepared repository of data designed to support decision making. The data comes from operational systems and external sources. To create the data warehouse, data are extracted from source systems, cleaned (e.g., to detect and correct errors), transformed (e.g., put into subject groups or summarized), and loaded into a data store (i.e., placed into a DW).

Benefits of using data warehouse in decision making
1. Improvements in turnaround time for data access and reporting generation.
2. Improved productivity of analytical staff due to availability of data.
3. Standardizing data across the organization so that there will be one view of information.
4. Merging data from various source systems to create a single information source system.
5. Business improvement resulting from analysis of warehouse data.
6. Encouraging and improving fact-based decision making.

The data warehouse supports OLAP, the functional and performance requirements of which are quite different from those of the OLTP applications traditionally supported by the operational databases.

OLTP covers most of the day to day operations of an organization such as purchasing, inventory, manufacturing, banking, payroll, registration and accounting. An OLTP system is customer oriented and is used for transaction and query processing by clerks, clients and information technology professionals. It manages the current data that, typically, are too detailed to be easily used for decision making. An OLTP system focuses mainly on the current data within an enterprise or department, without referring to historical data or data in different organizations. The access patterns of OLTP systems consist mainly of short, atomic transactions. Such systems require concurrency control and recovery mechanisms.

DW systems, on the other hand, are targeted for decision support. It serves users or knowledge workers in the role of data analysis and decision making. Such systems can organize and present data in various formats in order to accommodate the diverse needs of the different users.

An OLAP system is market-oriented and is used for data analysis by knowledge workers, including managers, executives, and analysts. These systems manage large amount of historical data, provides facilities for summarization and aggregation, and store and manages information at different levels of granularity. These features make the data easier to use in informed decision making. An OLAP system typically adopts either a star or a snowflake model and a subject-oriented database design. This system often spans multiple versions of a database schema, due to the evolutionary process of an organization. These systems also deal with information that originates from different organization, integrating information from multiple data stores. Because of their huge volume, OLAP data are stored on multiple storage media.

To facilitate complex analysis and visualization, the data in a warehouse is typically modeled multi dimensionally. For example, in a sales data warehouse, time of sales, sales district, salesperson, and product might be some of the dimensions of interest. Often these dimensions are hierarchical; time of sale may be organized as: day – month – quarter – year hierarchy, product as a product-category-industry hierarchy.

An operational database is designed and tuned from known tasks and workloads such as indexing and hashing using primary keys, searching for a particular record and optimizing ‘canned’ queries. On the other hand, data warehouse queries are often complex. They involve the computation of large amount of data at summarized levels and may require the use of special data organization, access and implementation methods based on multidimensional views.

DW might be implemented on Relational OLAP (ROLAP) servers. These are the intermediate servers that stand in between a relational back-end server and client front-end tools. They use a relational or extended-relational DBMS to store and manage warehouse data, and OLAP middleware to support missing pieces. They support extensions to SQL and special access and implementation methods to efficiently implement the multidimensional data model and operations. In contrast, Multidimensional OLAP (MOLAP) servers support multidimensional views of data through array-based multidimensional storage engines. They map multidimensional views directly to data cube array structures.

There is more to building and maintaining the data warehouse than selecting an OLAP server, defining a schema and some complex queries for the warehouse. Different architectural alternatives exist. Many organizations want to implement an integrated enterprise warehouse that collects information about all subjects spanning the whole organization. However, building an enterprise warehouse is long and complex process. Some organizations are settling for data marts instead, which are departmental subsets focused on selected subjects. These data marts enable faster roll out since they do not require enterprise-wide consensus.

II. ARCHITECTURE AND PROCESS DESIGN

![Figure: OLAP with data warehousing architecture](image-url)
A DW can be built using a top-down approach, a bottom-up approach, or a combination of both. The top-down approach starts with the overall design and planning. It is useful in cases where the technology is mature and well known, and where the business problems that must be solved are clear and well understood. The bottom-up approach starts with experiments and prototypes. This is useful in the early stage of business modeling and technology development. It allows an organization to move forward at considerably less expense and to evaluate the benefits of the technology before making significant commitments. In the combined approach, an organization can exploit the planned and strategic nature of the top-down approach while retaining the rapid implementation and opportunistic application of the bottom-up approach.

The design and construction of the DW consists of the following steps: planning, requirements study, problem analysis, warehouse design, data integration and testing, and finally deployment of the DW.

1. Choose a business process to model, for example, shipments, inventory, sales etc
2. Choose the grain of the business process. The grain is the Fundamental, atomic level of data to be represented in the fact table for this process.
3. Choose the dimensions that will apply to each fact table record.
4. Choose the measures that will populate each fact table record.

Once a data warehouse is designed and constructed, the initial deployment of the warehouse includes initial installation, roll out planning, training and orientation.

Designing and rolling out a data warehouse is a complex process, consisting of the following:

- Define the architecture, do capacity planning, and select the storage servers, database and OLAP servers, and tools.
- Integrate the servers, storage, and client tools.
- Design the warehouse schema and views.
- Define the physical warehouse organization, data placement, partitioning, and access methods.
- Connect the sources using gateways, ODBC drivers, or other wrappers.
- Design and implement scripts for data extraction, cleaning, transformation, load, and refresh.
- Populate the repository with the schema and view definitions, scripts, and other metadata.
- Design and implement end-user applications.
- Roll out the warehouse and applications.

### III. BACK-END TOOLS AND UTILITIES

Data warehouse systems use back-end tools and utilities to populate and refresh their data.

#### Data Cleaning:
Data cleaning routines attempt to fill in missing values, smooth out noise while identifying outliers, and correct inconsistencies in the data. Since a data warehouse is used for decision making, it is important that the data in the warehouse must be correct. Some examples where data cleaning becomes necessary are: inconsistent field length, inconsistent descriptions, inconsistent value assignments, missing entries and violation of integrity constraints.

#### Load:
After extracting, cleaning and transforming, data must be loaded into the warehouse. Additional preprocessing may still be required: checking integrity constraints; sorting; summarization; aggregation; and other computations to build the derived tables stored in the warehouse. In addition, load utility also allows the system administrator to monitor status, to cancel, to suspend and resume a load, and to restart after failure with no loss of data integrity. The load utilities for data warehouses have to deal with much larger data volumes than for operational databases.

#### Refresh:
Refreshing a warehouse consists in propagating updates on source data to correspondingly update the base data and derived data stored in the warehouse. There are two sets of issues to consider: when to refresh and how to refresh. Usually, the warehouse is refreshed periodically. The refresh policy is set by the warehouse administrator, depending on user needs and traffic, and may be different for different sources. Refresh techniques also depends on the characteristics of the source and capabilities of the database servers. Replication servers can be used to refresh a warehouse when the sources change.

### IV. MULTIDIMENSIONAL DATA MODEL

DW and OLAP tools are based on a multidimensional data model. This model views data in the form of a data cube. A data cube allows data to be modeled and viewed in multiple dimensions. Dimensions are perspectives or entities with respect to which an organization wants to keep records. Each dimension has a table associated with it, called a dimension table, which further describes the dimension.

A multidimensional data model is typically organized around a central theme. This theme is represented by a fact table. The fact table contains the names of the
facts, or measures, as well as keys to each of the related dimension tables.

![Figure: A logical view of representation a multidimensional data model in an OLAP cube](image)

V. FRONT-END TOOLS

The multidimensional data model grew out of the view of business data popularized by spreadsheet programs that are extensively used by business analysts. One of the popular operations that are supported by the multidimensional spreadsheet is pivoting. Pivot also called rotate, is a visualization operation that rotates the data axes in view in order to provide an alternative presentation of the data. Other operations are roll-up, drill-down, slice and dice. The roll-up operation performs the aggregation on a data cube, either by climbing up the concept hierarchy for a dimension or by dimension reduction. Drill-down is the reverse of the roll-up. It navigates from less detailed data to more detailed data. The slice operation performs a selection on one dimension of the cube. The dice operation performs a selection on two or more dimensions.

VI. DATA MINING

Data Mining is the extraction or “Mining” of knowledge from a large amount of data or data warehouse data mining should have been more appropriately named as “Knowledge Discovery from Data”.

Joseph P. Bigus, Defines the Data Mining in the following way. “The efficient discovery of valuable, non-obvious information from a large collection of data”.

Data Mining is the process of discovering interesting patterns and knowledge from large amounts of data. The data sources can include databases, data warehouses, the web, or other information repositories, or data that are streamed into the system dynamically.

To do this extraction data mining combines artificial intelligence, statistical analysis and database management systems to attempt to pull knowledge form stored data. Data mining is the process of applying intelligent methods to extract data patterns. This is done using the front-end tools. The spreadsheet is still the most compiling front-end application for OLAP. The challenges in supporting a query environment for OLAP can be crudely summarized as that of supporting spreadsheet operation effectively over large multi-gigabytes databases. To distinguish information extraction through data mining from that of a traditional database querying, the following main observation can be made. In a database application the queries issued are well defined to the level of what we want and the output is precise and is a subset of operational data. In data mining there is no standard query language and the queries are poorly defined. Thus the output is not precise and do not represent a subset of the database. Beside the data used not the operational data that represents the to day transactions. For instance during the process of building a data warehouse the operational data are summarized over different characteristics, such as borrowings during 3 months period. Queries can be of the type of “identify all borrowers who have similar interest” or “items a member would frequently borrow along with movies”, which is not a precise as the list of books borrowed by a member. The nature of the database and the query result in extracting non-subset of data. In supermarkets such relationships have already been identified using data mining. Thus related items such as “bread and milk’ or “beer and potato chips” would be kept together. Mobile companies decide on peak hours, rates and special packages based similar market research. Users can use data mining techniques on the data warehouse to extract different kinds of information which would eventually assist the decision making process of an organization. Such knowledge could only be discovered through sharing experiences of librarians or by capturing the knowledge through database and integrating them as done when building DW. DSS tools assist users in discovering knowledge.

VII. OLTP and OLAP

Previously, the job of on-line operational systems was to perform transaction and process the query. So, they are also termed as OLTP. Data warehouse systems serve users or knowledge workers in the role of data analysis and decision-making. Such systems can organize and present data in various formats in order to accommodate the diverse needs of the different users. These systems are called OLAP systems.

Pre-requisite of data warehousing and OLAP

Data warehousing developed, despite the presence of operational databases due to following reasons:

- An operational database is designed and tuned from known tasks and workloads, such as indexing using primary keys, searching for particular records and optimizing ‘canned queries’. As data warehouse
queries are often complex, they involve the computation of large groups of data at summarized levels and may require the use of special data organization, access and implementation methods based on multidimensional views.

- An operational database supports the concurrent processing of multiple transactions. Concurrency control and recovery mechanisms, such as locking and logging are required to ensure the consistency and robustness of transactions. While and OLAP query often needs read-only access of data records for summarization and aggregation. Concurrency control and recovery mechanisms, if applied for such OLAP operations, may jeopardize the execution of concurrent transactions.
- Decision support requires historical data, whereas operational databases do not typically maintain historical data. So, the data in operational databases, though abundant, is always far from complete for decision-making.
- Decision support needs consolidation (such as aggregation and summarization) of data from heterogeneous sources; and operational databases contain only detailed raw data, which serves as base for decisions which are outputs of the decision process one has to identify the problem first to arrive at proper decision.

VIII. DATABASE DESIGN

Most data warehouse use a star schema to represent the multidimensional data model. The database consists of a single fact table and a single table for each dimension. Each tuple in the fact table consists of a pointer to each of the dimension that provides its multidimensional coordinates and stores the numeric measures for that coordinates. Each dimension table consists of columns that correspond to attributes of the dimension.

The Snowflake schema is the variant of the star schema model, where some dimension tables are normalized, thereby further splitting the data into additional tables. The resulting schema graph forms the shape similar to a snowflake.

The major difference between the snowflake and star schema models is that the dimension table of the snowflake model may be kept in normalized form to reduce redundancies. Such a table is easy to maintain and saves storage space.

IX. METADATA REPOSITORY

Metadata are data about data. When used in a data warehouse, metadata are the data that define warehouse objects. Metadata are created for the data names and definitions of the given warehouse. Administrative metadata includes all of the information necessary for setting up and using a warehouse; description of the source databases; back-end and front-end tools. Business metadata includes business terms and definitions; ownership of the data. Operational metadata includes information that is created during the operation of the warehouse; monitoring information such as usage statistics, error reports, and audit trails. Metadata repository is used to store and manage all the metadata associated with the warehouse. The repository enables the sharing of the metadata among tools and processing for designing, setting up, using, operating and administering a warehouse.

Metadata play a very different role than other data warehouse, and are important for many reasons. For example, metadata are used as a directory to help the decision support system analyst to locate the contents of the data warehouse, as a guide to the mapping of the data when the data are transformed from the operational environment to the data warehouse environment. So, metadata should be stored and managed persistently.

X. CONCLUSION

Data warehouses have become base for effective tool for taking Managerial decisions so that tome lag is reduces and effective implementation takes place.
how this help in strategic decision making, which are required to be taken by top management of organization to run it effectively and successfully reaching to achieve the objective of business. Next, is construction the architecture of the data warehouse and the process of a data warehouse design by integrating data from multiple heterogeneous sources to support and /or adhoc queries, analytical reporting and decision making. Data warehouses provide online analytical processing (OLAP) tools for the interactive analysis of multidimensional data of varied granularities, which facilitates effective data mining. Data warehousing and online analytical processing (OLAP) are essential elements of decision support, which has increasingly become a focus of the database industry. OLTP is customer-oriented and is used for transaction and query processing by clerks, clients and information technology professionals. The job of earlier on-line operational systems was to perform transaction and query processing. Data warehouse systems serve users or knowledge workers in the role of data analysis and decision making. Next is designing of data warehouses, data mining, distinguished between data warehouse and other techniques (OLAP, OLTP etc).

Data warehouse do not contain the current information. However, data warehouse brings high performance to the integrated heterogeneous database system. It can store and integrate historical information and support complex multidimensional queries. As a result, data warehousing has become very popular in industry.

REFERENCES:

[3] Jiawei Han, Micheline Kamber and Jian Pei : Data Mining Concepts and Techniques.
[8] Pieter Adriaans and Dolf Zanting: Data Mining.
[9] Sam Anahory and Dennis Murray: Data Warehousing In the Real World.
AN ANALYTICAL SURVEY ON CLOUD COMPUTING
APPLICATIONS AND SECURITY ISSUES

ARATI KOTHARI

Department of Computer Science, CMJ University, Shillong, Meghalaya, India

Abstract - Cloud computing, a rapidly developing information technology has become a well-known the whole world. Cloud computing is Internet-based computing, whereby shared resources, software and information, are provided to computers and devices on-demand, like the electricity grid. Cloud computing is the product of the fusion of traditional computing technology and network technology like grid computing, distributed computing, parallel computing and so on. Companies, such as Amazon, Google, Microsoft and so on, accelerate their paces in developing Cloud Computing systems and enhancing their services to provide for a larger amount of users. However, security and privacy issues present a strong barrier for users to adapt into Cloud Computing systems. It is a new technology that satisfies a user’s requirement for computing resources like networks, storage, servers, services and applications, without physically acquiring them. It reduces the overhead of the organization of maintaining the large system but it has associated risks and threats also which include – security, data leakage, insecure interface and sharing of resources and inside attacks.

Keywords: Cloud Computing, on demand, pay per use, threats, data leakage, CSP

1. INTRODUCTION
Cloud Computing has more attention nowadays. The government agencies, enterprises, and more companies started to use Cloud Computing. The next revolution in IT and the big switch in IT is shifting the paradigm of classical computing to cloud environment. In Classical computing buying and owning the system software, hardware, application software to meet the peak needs. The system configuration, testing, verification and evaluation raise the demand of using resources. The cost of using the resource utilization is higher in order to meet the demands. In Cloud computing environment the storage of data and faster evaluation of CPU utilization and provides subscription of the data storage and pay per use of service as per the quality of services. Cloud computing is different from other technological trends but related with grid computing, utility computing and transparent computing. Grid computing is a base form of distributed computing whereby a super and virtual computing composed of clusters of network, loosely coupled computers acts to perform extreme tasks. Utility computing is the packaging of computing resources, such as computation and storage, as a metered service similar to a traditional public utility such as electricity, water, gas and telephone services. Transparent computing is a complex back-end services that are transparent to users which is simple and easy-to use provides front-end interface.

The term cloud has been used historically as a metaphor for the Internet. This usage was originally derived from its common depiction in network diagrams as an outline of a cloud, used to represent the transport of data across carrier backbones (which owned the cloud) to an endpoint location on the other side of the cloud. However, since the turn of the millennium, the concept has been revitalized. It was during this time of revitalization that the term cloud computing began to emerge in technology circles.

Definition of cloud computing:
• Cloud computing is a pay-per-use model for enabling available, convenient, on-demand network access to a shared pool of configurable computing resources like networks, servers, storage, applications, services etc, that can be rapidly provisioned and released with minimal management effort or service provider interaction.
• Cloud Computing is a paradigm that focuses on sharing data and computations over a scalable network of nodes. Examples of such nodes include end user computers, data centers, and Cloud Services. We term such a network of nodes as a Cloud.
• A Cloud is a type of parallel and distributed system consisting of a collection of interconnected and virtualized computers that are dynamically provisioned and presented as one or more unified computing resources based on service-level agreements established through negotiation between the service provider and consumers.”- Rajkumar Buyya, University of Melbourne.

The characteristics of a cloud server that includes the following:
• Ease of use-Includes simplified deployment and maintenance and portability.
• Instant availability- provides access to the server via a direct cabling that are connected
over the Local Area networks and the internet.

- Subset of functionality
- Privacy in nature- we own our data and control over the data.
- Overall controlling power of software resources.

Cloud Architecture

![Architecture of Cloud](image)

**Figure 1: Architecture of Cloud**

**Architecture**

- Application Services (services on demand)
  - Gmail, GoogleCalender
- Platform Services (resources on demand)
  - Google Appengine
- Infrastructure as services (physical assets as services)
  - IBM Blue house, Amazon EC2, Microsoft Azure Platform, Sun Parascale etc.

Multi-tenancy refers to the ability to run multiple customers on a single software instance installed on multiple servers. This is done to increase resource utilization by allowing load balancing among tenants, and to reduce operational complexity and cost in managing the software to deliver the service.

**Advantages of Cloud Computing.**

Cloud computing offers lots of advantages:

- Cost: As in the clouds the user needs not own the resources, it just needs to pay as per the usage in terms of time, storage and services. This feature reduces the cost of owning the infrastructure.
- Performance: the performance is improved because the cloud is not a single computer but a large network of powerful computers resulting in high processing power.
- Freedom from up gradation and maintenance: the cloud infrastructure is maintained and upgraded by the cloud service provider.
- Scalability: The user is can request to increase the resources if the area of application grows or new functionality is added. On the other hand if requirement shrinks the user can request to reduce the resources as well.
- Speedy Implementation: Time of Implementation of cloud for an application may be in days or sometimes in hours. You just need a valid credit card and need to fulfill some online registration formalities.
- Mobility: We don’t need to carry our personal computer, because we can access our documents anytime anywhere.
- Increase Storage Capacity: In Cloud computing we have extreme resources for storing data because our storage consists of many bases in the Cloud. Another thing about storing data in the Cloud is that, because of our data in the Cloud can automatically duplicated, they will be more safety.

**Disadvantages of Cloud Computing**

There are certain security threats and issues of implementing Cloud computing:

- Data Loss: Customers are responsible for the security of their data, thus in any case if data is lost the customer is in deep trouble.
- Account Hijacking: Since No native APIs are used for login and anyone can easily register himself as cloud service user chances of hijacking ones account are very high.
- Control over the process – in cloud computing the user have very less or no control over the services.
- Insider attacks by Cloud Service Provider: It may possible that a fraudulent employee may do the fishing and steal the data.
- Legal aspects: In case of Data loss the user may suffer if there is no Service Level Agreement (SLA), the loss will be of user, because he is not able to put claims against the cloud service provider.

II. SERVICES IN CLOUD

There are three types of cloud providers that you can subscribe to: Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS). These three types differ in the amount of control that you have over your information, and conversely, how much you can expect your provider to do for you. Briefly, here is what you can expect from each type.

1. **Software as a Service** - A SaaS provider gives subscribers access to both resources and applications. SaaS makes it unnecessary for you to have a physical copy of software to install on your devices. SaaS also makes it easier to have the same software on all of your devices at once by accessing it on the cloud. In a SaaS
agreement, you have the least control over the cloud.

2. **Platform as a Service** - A PaaS system goes a level above the Software as a Service setup. A PaaS provider gives subscribers access to the components that they require to develop and operate applications over the internet.

3. **Infrastructure as a Service** - An IaaS agreement, as the name states, deals primarily with computational infrastructure. In an IaaS agreement, the subscriber completely outsources the storage and resources, such as hardware and software that they need.

As you go down the list from number one to number three, the subscriber gains more control over what they can do within the space of the cloud. The cloud provider has less control in an IaaS system than with an SaaS agreement.

### III. TYPES OF CLOUDS

There are different types of clouds that you can subscribe to depending on your needs. As a home user or small business owner, you will most likely use public cloud services.

1. **Public Cloud** - A public cloud can be accessed by any subscriber with an internet connection and access to the cloud space.

2. **Private Cloud** - A private cloud is established for a specific group or organization and limits access to just that group.

3. **Community Cloud** - A community cloud is shared among two or more organizations that have similar cloud requirements.

4. **Hybrid Cloud** - A hybrid cloud is essentially a combination of at least two clouds, where the clouds included are a mixture of public, private, or community.

### IV. CLOUD SECURITY ISSUES

Security issues come under many guises both technical and socio-technical in origin. To cover all the security issues possible within the cloud, and in-depth, would not be possible. Existing efforts look to provide taxonomy over the issues seen. The Cloud Security Alliance is a non-profit organization that seeks to promote the best practices for providing security assurance within the cloud computing landscape. The Cloud Security Alliance identify seven threats to cloud computing that can be interpreted as a classification of security issues found within the cloud. They are:

1. **Abuse and Nefarious Use of Cloud Computing**

    Legitimate CSPs can be abused for nefarious purposes, supporting criminal or other untoward activities towards consumers. The emphasis is that legitimate services are used with malicious intent in mind. Other issues seen include the provision of purposefully insecure services used for data capture. Commonly asked information includes the consumers name, email/postal address, D.O.B or even credit card details. Users are essentially being goaded to part with more information than required as a prerequisite for service use. Malicious entities can then use this information for nefarious purposes, most notably identity theft. Even if the entities are not malicious the disclosure of information to the CSP can also be said to be an abuse of service by the CSP themselves. CSPs may collect this information, or any other information provided at later stages, and markets this information to third parties for data mining purposes. Another security issue found is the provision of legitimate maliciously-oriented services. This service provision also known as (In) Security-as-a-Service, offers users the same security guarantees as existing services yet their use is malicious. By using a cloud based service it is claimed that a process that would take around five plus days now takes on average twenty minutes.

2. **Insecure Interfaces and Application Programming Interfaces**

    Data placed in the Cloud will be accessed through Application Programming Interfaces (APIs) and other interfaces. Malfunctions and errors in the interface software, and also the software used to run the Cloud, can lead to the unwanted exposure of user’s data and impugn upon the data's integrity. A HTTP server can allow an attacker to gain complete control over the web server. Data exposure can also occur when a software malfunction affects the access policies governing user’s data. This has been seen in several Cloud based services in which a software malfunction resulted in which a user's privacy settings shall be overwritten and the user data is exposed to non-
authorized entities. Threats can also exist as a result of poorly designed or implemented security measures. If these measures can be bypassed, or are non-existent, the software can be easily abused by malicious entities. Regardless of the threat origin, APIs and other interfaces need to be made secure against accidental and malicious attempts to circumvent the APIs and their security measures.

3. Malicious Insiders
Although a CSP can be seen as being honest their employees may not be. A malicious insider may be an employee of the CSP who abuses their position for information gain or for other nefarious purposes. Regardless, of the employee’s motivation the worrying aspect is that a surreptitious employee will have access to consumer’s data for legitimate purposes but will abuse this power for their own means. Another more subtle form of the malicious insider problem is through PaaS based services. If the service provider offers a platform that allows developers the ability to interact with user’s data i.e. Social Networking Applications, users may unknowingly allow these developers access to all their data. Use of this platform may be unchecked. For example, it is well known on the Social Networking Web Sites. Platform that once a user adds an application the application will have the ability to access the entire user’s information, if allowed to do so, regardless of the applications function. Even if the application developers are not malicious this does not mean that the application cannot be hacked.

4. Shared Technology Issues
A more interesting form of confidentiality issue relates to the construction of a cloud and the services themselves.

4.1. Virtualization Issues
The underlying virtualization architecture allows IaaS service providers the ability to host several machine images on a single server. Practical attacks on such services are: First, the authors showed that they could map the internal structure of the cloud, allowing them to determine if two virtual machines were co-resident with each other i.e. were running on the same physical machine. Secondly, they demonstrated that they were able to, purposefully; add a virtual machine to the cloud so that it was co-resident with another machine. Finally, the authors were able to show that once a machine was co-resident, they would be able to launch several attacks that would allow them to learn information regarding CPU cache use, network traffic rates and keystroke timings.

4.2. Service Aggregation
Aggregated services offer services based upon the functionality offered by existing services. Often aggregated services offer the combined functionality of existing services allowing for rapid service construction. However, service aggregation presents consumers with several interesting problems. Data is now being shared across multiple service providers whose privacy policies will also subject to change. Under whose privacy policy is the data governed by, how to combine the two policies? Furthermore, service aggregation can occur in an ad-hoc and rapid manner implying that less stringent controls could have been applied to the protection of data, increasing the likelihood of a problem.

5. Data Loss or Leakage
Although insecure APIs can lead to data loss or the unwanted exposure of information, consumers can also lose their information through other means.

5.1. Availability Issues
Availability issues are when user’s data is made inaccessible to the consumer. The data has been made unavailable. Such a lack of availability can be a result of access privilege revocation, data deletion or restricting physical access to the data itself. Availability issues can be attributed to an attacker using ooding based attacks. A monetary effect can also be seen with availability issues. Monetary issues affect not only the consumer but also the CSP. Flooding attacks will have an effect upon resource utilization and will result in increases to power consumption, network usage and hardware maintenance. Ultimately this will also increase the amount of money the consumer will be charged for resource usage. Moreover, these monetary increases will also increase the operational expenditure of the service provider. Fault tolerance protocols are used to combat the issue of node failure within the Cloud. Fault tolerance protocols replicate data across machines, and data centers, ensuring that if part of the cloud does fail a version of the data will still be available to the user. Part of the cloud will be redundant though for good measure. Data replication also requires that the data state of the data be synchronized across multiple nodes. However, poorly designed protocols can also lead to availability issues. Furthermore, the existence of multiple copies of data can also introduce confidentiality problems. The increased number of data instances also increases the likelihood that an attacker will be able to access the data.

5.2. Data Leakage
Another form of data leakage stems from the disclosure of information that, though hidden, is deduced from freely available information. When users interact with a service they can leave a public trail, be it from status/update messages or through
new postings. Unwelcome linkage occurs when new information is discerned about an individual through analysis of the individuals public trail i.e. links. This unwelcome linkage could be accidental or the result of the individual not covering their tracks. Social graph merging is similar to unwelcome linkage however the links formed occur through the aggregation of social graphs. A social graph is a graph describing a person's social information such as friends, groups and interests.

Through combination of a person’s social graphs from separate social networking sites, or the social graphs of people from the same social network site, new information can be deduced.

6. Account or Service Hijacking

When communicating with the CSP malicious entities may seek to affect the integrity and authenticity of the user's communication with the CSP and vice versa. There are several ways in which the integrity and authenticity of a user’s session can be impugned. Browser based interfaces and authentication are used by consumers to establish a session with their service provider. A malicious entity can attempt to capture or hijack this session or steal the user’s credentials to access or influence the user’s data, from within the browser. Most browsers operate on a Same Origin Policy where client scripts are allowed to access resources if they share the same origin. However, attacks such as Cross Site Scripting and Cross Site Request Forgery, as well as DNS Poisoning can be used to undermine this security feature. Other issues found include the manipulation of the data being sent within the sessions. The effects of breaking session integrity are two-fold, for one the attacker will be able to steal the identity of their victim, and secondly impugn the reputation of the victim through falsified data. Such man-in-the-middle attacks will have lasting repercussions such as violation of the services terms of use, or criminal. For instance, the hacking of many a celebrity on social networking account.

7. Unknown Risk Profile

Risk Management is a business process that users can use to identify and mitigate threats. It allows users to determine their current stance towards the security of their data. Auditing information such as software version, code updates current security practices, intrusion attempt are used as a basis for determining this stance. CSPs may not be so forthcoming with this information. Consumers, when adopting a service, must also accept the Terms and Conditions (including privacy policy) of the service, together with any Service Level Agreements made. Consumers and providers need to comply with existing laws and regulations. However the degree to which service providers adhere to current security practices and legislation, or implement them may not be clear. This leaves the consumers with an unknown risk profile. Users are unable to determine the risk to their data as they do not have sufficient information to do so.

V. CONCLUSIONS

Cloud computing is still struggling in its accuracy, with positive and negative comments made on its possible implementation for a large-sized enterprise. IT technicians are spearheading the challenges. Cloud computing provides many options for the everyday computer user as well as large and small businesses. It opens up the world of computing to a broader range of uses and increases the ease of use by giving access through any internet connection. However, with this increased ease also come drawbacks. You have less control over who has access to your information and little to no knowledge of where it is stored. You also must be aware of the security risks of having data stored on the cloud. The cloud is a big target for malicious individuals and may have disadvantages because it can be accessed through an unsecured internet connection.

If you are considering using the cloud, be certain that you identify what information you will be putting out in the cloud, which will have access to that information, and what you will need to make sure it is protected. Additionally, know your options in terms of what type of cloud will be best for your needs, what type of provider will be most useful to you, and what the reputation and responsibilities of the providers you are considering are before you sign up.

REFERENCES

http://www.infoworld.com/d/securitycentral/gartnerseve
n-cloud-computing-security-risks-853.


PERFORMANCE EVALUATION OF PVM BASED PARALLEL APPLICATIONS USING A CLUSTER ORIENTED PARALLEL COMPUTING ARCHITECTURE

SAMPATH S1, BHARAT BHUSHAN SAGAR2 & NANJESH B R3

1Shri Venkateshwara University, Gajraula, Uttarpradesh, INDIA
2Department of Information Technology, Ajay Kumar Garg Engineering College Ghaziabad, Uttarpradesh, INDIA
3Department of Computer Science and Engineering, Adichunchanagiri Institute of Technology Chikmagalur, Karnataka, INDIA

Abstract—Parallel computing operates on the principle that large problems can often be divided into smaller ones, which are then solved concurrently to save time (wall clock time) by taking advantage of non-local resources and overcoming memory constraints. The main aim is to form a cluster oriented parallel computing architecture for PVM based applications which demonstrates the performance gain and losses achieved through parallel processing using PVM. This can be realized by implementing the parallel applications like solving matrix multiplication problem, using PVM. The architecture for demonstrating PVM based parallel applications works on the Master-Slave computing paradigm. The master will monitor the progress and be able to report the time taken to solve the problem, taking into account the time spent in breaking the problem into sub-tasks and combining the results along with the communication delays. The slaves are capable of accepting sub problems from the master and finding the solution and sending back to the master. We aim to evaluate these statistics of parallel execution and do comparison with the time taken to solve the same problem in serial execution to demonstrate communication overhead involved in parallel computation. The results with runs on different number of nodes are compared to evaluate the efficiency of PVM based parallel applications. We also show the performance dependency of parallel and serial computation, on RAM.

Keywords—Parallel Execution, Cluster Computing, Symmetric Multi-Processor (SMP), PVM (Parallel Virtual Machine), RAM (Random Access Memory).

1. INTRODUCTION

Parallel processing refers to the concept of speeding up the execution of a program by dividing the program into multiple fragments that can execute simultaneously, each on its own processor. This paper deals how to handle Matrix Multiplication problem that can be split into sub-problems and each sub-problem can be solved simultaneously. With computers being networked today, it has become possible to share resources like files, printers, scanners, fax machines, email servers, etc. One such resource that can be shared but is generally not, is the CPU. Today's processors are highly advanced and very fast, capable of thousands of operations per second. If this computing power is used collaboratively to solve bigger problems, the time taken to solve the problem can reduce drastically.

A. Existing Frameworks

1) PVM 1.0: PVM that have been released from the first one in February 1991. PVM 1.0 cleaned up the specification and implementation to improve robustness and portability [2].

2) PVM 2.X Versions: PVM 2.1 provided with process-process messages switched to XDR to improve portability of source in heterogeneous environments and simple console interpreter added to master pvmid. Later versions belonging to PVM 2.x provided with more and more useful functionalities such as pvmid-pvmid message format switched to XDR, get and put functions vectorized to improve performance, broadcast function deprecated, improved password-less startup via rsh/rcmd etc [2].

3) PVM 3.X Versions: These allow scalability to hundreds of hosts, allow portability to multiprocessors / operating systems other than Unix, allows dynamic reconfiguration of the virtual machine, allows fault tolerance, includes dynamic process groups, provide the option to send data using a single call [3].

4) PVM 3.4.6: Includes both Windows and UNIX versions and improved use on Beowulf clusters. Also includes the latest patches for working with the latest versions of Linux (like fedora 14), Sun, and SGI systems New features in PVM 3.4.x include communication contexts, message handlers, persistent messages. In our project, we are using PVM 3.4.6 for providing parallel environment using PVM [4].

B. Proposed System

This paper deals with the implementation of parallel application, matrix multiplication under PVM using PVM 3.4.6 for communication between the cores and for the computation. Because they are very much suitable to implement in LINUX systems.
II. RELATED WORKS

Traditionally, multiple processors were provided within a specially designed "parallel computer"; along these lines, Linux now supports SMP Pentium systems in which multiple processors share a single memory and bus interface within a single computer. It is also possible for a group of computers (for example, a group of PCs each running Linux) to be interconnected by a network to form a parallel processing cluster [1]. V.S Sunderam, G.A Geist, J Dongarra, R Manchek (1994) [6] describe the architecture of PVM system, and discuss its computing model, the programming interface it supports, auxiliary facilities for process groups and MPP support, and some of the internal implementation techniques employed. Amit Chhabra, Gurvinder Singh (2010) [7] proposed Cluster based parallel computing framework which is based on the Master-Slave computing paradigm and it emulates the parallel computing environment. Muhammad Ali Ismail, Dr. S. H. Mirza, Dr. Talat Altaf (2011) [8] performed the Concurrent Matrix Multiplication on Multi-Core Processors. Kamalrulnzam Abu Bakar, Zaitul Milir Izawati Zal nuddln (2006) [9] made the performance comparison of PVM and RPC. The comparison is done by evaluating their performances through two experiments namely one is a broadcast operation and the other are two benchmark applications, which employ prime number calculation and matrix multiplication. We aim to present a architecture over which PVM based parallel application runs and which demonstrates the performance gain and losses achieved through parallel processing. And also demonstrates the performance dependency of parallel applications on RAM.

III. SYSTEM REQUIREMENT

A. Hardware Requirements
- Processor: Pentium D (3 G Hz)
- Two RAM: 256MB and 1GB
- Hard Disk Free Space: 5 GB
- Network: TCP/IP LAN using switches or hubs

B. Software Requirements
- Operating System: Linux
- Version: Fedora Core 14
- Compiler: GCC

III. SYSTEM DESIGN

A. System Analysis
The system is to be designed such that it demonstrates the performance dependency of parallel and serial execution on RAM and also it demonstrates the following:
- How a client can submit the entire problem to a master and collects the solution back from it without bothering about how it has been solved.
- How the master detects the available slaves on the network, and how it detects the system load on that machine to determine whether it is worth sending a task to that particular client.
- How a problem can be submitted to the slaves.
- How the solutions of the given problem can be retrieved from the slave.
- How the slaves solve the given problem.

The design was made modular i.e. the software is logically partitioned into components that perform specific functions and sub-functions.

1) Master is designed such that it has functionality to manage connection and communication with the slave, It scans and identifies all the cores or slaves available on the node here it is only one slave to be identified. It then assigns the processor ranks to identify the cores. The master assigns the problem to slave. It also has to accept the results sent back by the slave after they finish the computation of the sub-tasks assigned to them. Then the received result has to be assembled in the right order to obtain the solution for the main problem.

2) Slave: This is designed to have the functionality to read the problem (in case of single slave)/sub-problem sent by the master, evaluate the problem (in case of single slave)/sub-problem and send the result back to the master.

B. Cluster Based Parallel Computing architecture
The main problem is taken by the master core and assigns the task into slave cores. Each slave core sends back the solutions of the assigned sub problem. The working principle involved in this architecture is shown in Fig.1 and Fig.2 shows the cluster based parallel computing architecture.

![Fig 1. Operations involved in cluster based parallel computing architecture](image-url)
Download the software package from http://www.netlib.org/pvg3
Unpacking: Command: $tar zxvf pvm3.4.6.tgz
Opened the .bhrc file through terminal using vi editor and set the following lines in the file [3] and closed the file. The .bhrc is a hidden file of course and can be done as:
$home
$ls -a
$vi .bhrc
Going to the insert mode add the following lines as:
PVM_ROOT=$HOME/pvm3
PVM_DPATH= PVM_ROOT/lib/pvmd
Export PVM_ROOT PVM_DPATH
Going to pvm3 directory ($cd pvm3) type make
($make). This would make pvm(the PVM console), pvmd3(the pvm daemon), libpvm3.a(PVM C/C++ library), libfpvm3.a (PVM Fortran library) and libgpv3.a (PVM group library). All these files would be placed in the $/pvm3/lib/LINUX and pvmgs (PVM group server) would be placed in $/pvm3/bin/LINUX. Open .rhosts file in the home directory ($ vi .rhosts) and added the name of the node (computer) in the cluster of four computers.
Set the following environment variables in .bashrc file as:
export PVM_ARCH='$PVM_ROOT/lib/pvmgetarch'
export PVM_ROOT= $HOME/pvm3
export PATH =$PATH:$PVM_ROOT/lib
export PATH=$PATH:$PVM_ROOT/lib/$PVM_ARCH
Going to the pvm3 directory and again executed make command. If a prompt pvm> is got means pvm is successfully installed.

V. IMPLEMENTATION

Implementation is the most crucial stage in achieving a successful parallel system. The problem to be solved has to be parallelized so that computation time is reduced. The framework consists of a client, a master core, capable of handling requests from the client, and slave, capable of accepting problems from the master and sending the solution back. The master and the slave communicate with each other using PVM3.4.6 under PVM. The problem has to be divided such that the communication between the server and the client is minimum. The total computational time to solve the problem completely is effected by the communication time between the nodes.

A. Parallel Matrix Multiplication Design

In the algorithm which we have implemented is for solving matrix multiplication problem on several nodes it may be for only one or more slaves. It divides the matrix into set of rows and sends it to the slaves rather than sending one row at a time [6]. The slaves compute the entire set of rows that they have received and send it back to the server in one send operation. Hence, we need to implement parallel systems consisting of set of independent desktop PCs interconnected by fast LAN cooperatively working together as a single integrated set of independent desktop PCs interconnected by fast LAN cooperatively working together as a single integrated computing resource so as to provide higher availability, reliability and scalability. But to show the performance dependency on RAM we are considering only single node with two cores, one acts as master and other as slave. So there will be no division of problem, instead entire problem is submitted to the single available slave. Rest of the work is carried out with the multiple nodes. Consider two matrix, matrix A and B. The flow of multiplication of matrix A and B takes place as shown in Fig. 3. The operations involved in dividing first matrix into set of rows and multiplying each set with entire second matrix giving resultant matrix is shown in Fig. 4.
Fig. 3. Flow diagram for solving matrix multiplication problem on several node

Fig. 4. Parallel matrix multiplication design

Fig. 5. Output of 1000 * 1000 matrix in parallel execution using four nodes

TABLE I

<table>
<thead>
<tr>
<th>Type of Execution</th>
<th>RAM size</th>
<th>100 * 100</th>
<th>500 * 500</th>
<th>1000 * 1000</th>
<th>1500 * 1500</th>
<th>2000 * 2000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serial</td>
<td>Two</td>
<td>0.0159 seconds</td>
<td>1.2744 seconds</td>
<td>9.8071 seconds</td>
<td>32.9938 seconds</td>
<td>78.1319 seconds</td>
</tr>
<tr>
<td></td>
<td>Three</td>
<td>0.0157 seconds</td>
<td>1.2713 seconds</td>
<td>9.8112 seconds</td>
<td>32.9798 seconds</td>
<td>78.1526 seconds</td>
</tr>
<tr>
<td>Parallel with MPI</td>
<td>Two</td>
<td>0.2521 seconds</td>
<td>5.4241 seconds</td>
<td>24.0342 seconds</td>
<td>78.2816 seconds</td>
<td>345.4321 seconds</td>
</tr>
<tr>
<td></td>
<td>Three</td>
<td>0.1121 seconds</td>
<td>3.4212 seconds</td>
<td>20.1462 seconds</td>
<td>70.2437 seconds</td>
<td>166.5418 seconds</td>
</tr>
</tbody>
</table>
VI. TESTING

There is a need of testing whether the resultant matrix is correct or not. Hence testing is one of the important step that must be performed, which is explained in the following section. Separate testing is made for smaller order matrices and larger order matrices.

A. Testing of Small Order Matrices

We have tested the output of our system by verifying the output using a online matrix calculator. We have repeated this test for different smaller sizes of matrix. The tool used by us is Blue Bit online matrix multiplication calculator.

B. Testing of higher order matrices

Blue Bit online matrix multiplication calculator is limited to matrix of order 32 as the testing tool does not support the matrix multiplication of higher order. To verify our result for higher order matrix we have assigned 1 to all the elements of the matrix. The elements of the output matrix will be equal to the order of the matrix. An example is shown in Fig 5.

VII. RESULTS AND ANALYSIS

We have analyzed the performance of parallel method against traditional serial method. The results are tabulated and compared. We calculated the time for solving the matrix multiplication problem using both serial and a case of parallel algorithm using PVM. Analysis can be made under four different cases.

Case 1: Single node analysis to show performance dependency on RAM. Table I shows that performance of serial execution almost remains same even after the increase in RAM size. There are negligible computation time variations for increase in RAM size. This is because the Serial execution is performed by the cores itself with negligible RAM usage and also due to the no communication involved between cores. Hence it is independent of RAM. We can also conclude that performance of parallel execution in both the cases MPI and PVM drastically increases when there is increase in RAM size. It shows drastic decrease in computation time with the increase in RAM. Because parallel execution often uses RAM for the communication between cores and also it involves lot of send and receive operations and temporarily storing the result of problem assigned to cores. We can analyze that higher the size of matrices the time difference is very high in the table, because higher the matrix size, more will be sends and receives resulting in the need of higher utilization of RAM. So for smaller RAM the computation time will be more and larger the RAM size computation time will be less in parallel execution finally resulting in better performance. However by seeing the time results for higher order matrices such as 1000*1000, 2000*2000, there is a large reduction in computation time for PVM when there is increase in RAM size. Hence we can conclude that PVM based parallel applications is more dependent on RAM size.

Case 2: Single node analysis to show communication overhead involved in parallel computations. Generally we can say parallel execution is faster than the serial execution but the results of serial execution with 1000MB RAM and parallel execution using PVM with 1000MB RAM shown in the Table I depicts that serial execution is faster than parallel execution in a single node having two cores. We can analyze that higher the size of matrices temporarily storing the result of problem assigned to cores. Hence it is independent of RAM. We can also conclude that performance of parallel execution is faster than parallel execution in a single node having two cores. For different sizes of matrices. This is due to the communication overhead involved in the parallel execution but this can be overcome by increasing the number of nodes but at present it is out of scope of our work and can be done as future work. Overheads that are considered are the connection time required to connect to slave, time taken to send the problem along with inputs to slave time taken to retrieve the solutions from the client, time taken to assimilate the results obtained.

Case 3: Multiple nodes analysis with smaller order matrices (computation time < communication time). Table II shows the time

![Table II: Performance of MPI and PVM for Smaller Order Matrices](image)

<table>
<thead>
<tr>
<th>Type Of Execution</th>
<th>Number of nodes</th>
<th>100 * 100</th>
<th>200 * 200</th>
<th>300 * 300</th>
<th>400 * 400</th>
<th>500 * 500</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parallel with PVM at 256 MB RAM</td>
<td>Two</td>
<td>0.4852 seconds</td>
<td>2.1346 seconds</td>
<td>3.1402 seconds</td>
<td>4.4637 seconds</td>
<td>11.3120 seconds</td>
</tr>
<tr>
<td></td>
<td>Three</td>
<td>0.9143 seconds</td>
<td>2.6711 seconds</td>
<td>3.7841 seconds</td>
<td>4.8761 seconds</td>
<td>11.9160 seconds</td>
</tr>
</tbody>
</table>

![Table III: Performance of MPI and PVM for Higher Order Matrices](image)

<table>
<thead>
<tr>
<th>Type of execution</th>
<th>Number of nodes</th>
<th>1000 * 1000</th>
<th>1500 * 1500</th>
<th>2000 * 2000</th>
<th>2500 * 2500</th>
<th>3000 * 3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parallel with PVM at 256 MB RAM</td>
<td>Two</td>
<td>32.5426 seconds</td>
<td>232.7810 seconds</td>
<td>464.6324 seconds</td>
<td>625.6328 seconds</td>
<td>843.7486 seconds</td>
</tr>
<tr>
<td></td>
<td>Three</td>
<td>21.3216 seconds</td>
<td>143.4632 seconds</td>
<td>216.4327 seconds</td>
<td>355.6449 seconds</td>
<td>497.1042 seconds</td>
</tr>
</tbody>
</table>
taken to solve the problem wholly is more when the
number of nodes is more for smaller matrix. Because
the problem has to be communicated among all the
slave cores hence the communication time is larger
than the computation time. So the 2 nodes can
compute it and assemble it faster than a 3 node or a 4
node system.

**Case 4: Multiple nodes analysis with higher
order matrices (computation time >
communication time).** Table III shows that for matrix
of higher order the performance of the system
increases phenomenally with increase in number of
nodes. As the size of the matrix increases the
computation time also increases. The computation
time is so large that the communication time is
negligible compared to it.

**VIII. CONCLUSION**

We presented a model that demonstrates the
performance gain and losses achieved through parallel
processing. Matrix multiplication problem is solved
serially and also in parallel under PVM. We
demonstrated the evaluation of the performance
dependency of PVM based parallel applications and
serial execution on RAM under different sizes of
RAM. Serial execution is faster for smaller matrix
because of the communication and connection
overheads in parallel execution. The performance of
parallel execution is far greater compared to serial
execution when the size of the matrix is large. The
total time taken to compute the result decreases
dramatically when the number of nodes increases.

**REFERENCES**

Framework Support Adaptive Applications in
Distributed/Parallel Computing” Department of Computer
Science, University College London United Kingdom.

[2] History of PVM versions:

National Lab., Oak Ridge, Tennessee.


Using PVM “, International Journal of Advanced Research in
Computer Science and Software Engineering (IJARCSSE),

Concurrent Computing System Evolution, Experiences and
Trends”, Parallel Computing - Special issue: message passing
interfaces archive Volume 20 Issue 4 Pages 531-545, April
1994.

Computing Framework (CBPCF) for Performance
Evaluation of Parallel Applications”, International Journal of

[8] Muhammad Ali Ismail, Dr. S. H. Mirza, Dr. Talat Altaf,
“Concurrent Matrix Multiplication on Multi Core Processors”,
International Journal of Computer Science and Security
(IJCSS), Volume (5) : Issue (2) : 2011.

[9] Kamalrulnzam Abu Bakar, Zaitul Mlir lizawati Zal nuddln,
“Parallel Virtual Machine versus Remote Procedure Calls: A
Performance Comparison”, Jilid IS.Bit. I, Jurnal Teknologi
Maklumat, Jun 2006.
E-LEARNING: A MODERN TEACHING-LEARNING ENVIRONMENT

MALATESH S. AKKUR1, S.R. MANGALWEDE2 & D.H. RAO3
1Dept. of Electronics, JAIN UNIVERSITY JC Road Bangalore, India.
2Department of Computer Science, JAIN ENGINEERING COLLEGE Belgaum, India.
3Principal and Director, JAIN ENGINEERING COLLEGE Belgaum, India.

Abstract-The e-Learning uses computer based learning, training and teaching materials, online conferencing, discussion boards, e-mail, computer-aided assessment, and other related methods. There are tremendous gains from practicing e-Learning through e-Learning tools, provided it is done carefully and systematically. The e-Learning when used effectively is not simply an add-on, but impacts all aspects of the teaching-learning environment. In this paper we discuss the need, aspects, pros & cons of e-Learning.

Key Terms: e-Learning, online conferencing, discussion boards, e-mail, computer-aided assessment, e-Learning tools, aspects of e-Learning.

1. INTRODUCTION

Almost everyone, some way connected to education, would have come across the term e-Learning. A simple google search on e-Learning would bring up many thousands of pages. One can find dozens of books on the topic. But a closer inspection of these materials and people’s perception of the meaning of the term e-Learning shows a wide range of variation. e-Learning does indeed have many dimensions and approaches, with associated spheres of influence. One size does not fit all. Often there is a tendency to copy models being tried out by somebody else, without bothering to carryout a detailed evaluation of that approach particularly in the context of the requirements and constraints present in one’s own environment. Such perception variation can be seen not just at the management level, but also among teachers and students. Hence e-Learning is very helpful to meet the individual academic requirements. The remaining paper is organized as follows: In section 2 we discuss about the definitions of online learning and e-Learning and analyze how the e-Learning is defferent from online education and distance education. In section 3 we focus on the need for e-Learning. In section 4 we explain the various aspects involved in the e-Learning. In section 5 we point out some limitations of e-learning. Finally in section 6, we conclude the objective of this paper.

2. DEFINITIONS OF ONLINE EDUCATION AND E-LEARNING:

Online Education: The different terms for online education are: virtual education, Internet-based education, web-based education, education via computer-mediated communication, etc. The definition of online education can be developed from the definition of Keegan:

“Distance education is a form of education characterized by:
- The quasi-permanent separation of teacher and learner throughout the length of the learning process (this distinguishes it from conventional face-to-face education);
- The influence of an educational organization both in the planning and preparation of learning materials and in the provision of student support services (this distinguishes it from private study and teach yourself programmes);
- The use of technical media – print, audio, video or computer – to unite teacher and learner and carry the content of the course;
- The provision of two-way communication so that the student may benefit from or even initiate dialogue (this distinguishes it from other uses of technology in education); and
- The quasi-permanent absence of the learning group throughout the length of the learning process so that people are usually taught as individuals rather than in groups, with the possibility of occasional meetings, either face-to-face or by electronic means, for both didactic and socialization purposes.” [17]

If online education is considered to be a subset of distance education we may define online education by accepting Keegan’s definition with a modification as:

“Online Education is the usage of the computer and internet technology by the teachers and learners to carry the contents of the course and to have the provision of two-way communication for the optimum benefit of teaching-learning process”.

Most of the characteristic of online education would exclude Keegan’s last point, as collaborative learning, where students may communicate throughout the length of the learning process is seen as one of the greatest advantages of online learning relative to previous “generations” of distance education [10]. On the other hand, there is good reason to stress that
most adult students need to organise their studies according to demands of work, social life and family responsibilities. These needs must be balanced against a possible didactic idea of co-operative learning. Thus, the flexibility of the institution in adapting course requirements so that students may organise their learning independent of a study group is important for many online students [18]. This does not at all exclude learning methods exploiting the advantages of being part of a group or learning community.

‘Distance education’ and ‘Distance learning’ as defined by Keegan are well-established concepts. The ‘Distance learner’ is a person who, for some reason, will not or cannot take part in educational programmes that require presence at certain times or places [17].

‘e-Learning’
The terms ‘e-Learning’ and ‘m-Learning’ have entered the scene and are replacing the terms ‘Distance education’ and ‘Distance learning’. These terms are taking the control of the teaching and learning community. Most of e-Learning programmes seem to be extremely costly to develop and implement.

e-Learning is defined in various ways by different researchers, teaching and learning community. Dichanz, a professor of education and the German FernUniversität, ends his critical analysis of the term, e-Learning with the following definition:

“e-Learning is the collection of teaching – and information packages – in further education which is available at any time and any place and is delivered to learners electronically. They contain units of information, self-testing batteries and tests, which allow a quick self-evaluation for quick placement. e-Learning offers more lower level learning goals. Higher order goals like understanding, reasoning and (moral) judging are more difficult to achieve. They require an individualised interactive discourse and can hardly be planned”[3]

We may not completely agree with the Dichantz, but we realize that higher level learning objectives may not be completely incorporated in e-Learning.

Unfortunately, the term e-Learning is often used as a more generic term and as a synonym for online education. Kaplan-Leiserson has developed an online e-Learning glossary. It provides the definition as:

‘e-Learning covers a wide set of applications and processes, such as Web-based learning, computer-based learning, virtual classrooms, and digital collaboration. It includes the delivery of content via Internet, intranet/extranet (LAN/WAN), audio-and videotape, satellite broadcast, interactive TV, and CD-ROM.

In the glossary of e-Learningeuropa.info, e-Learning is defined as: “The use of new multimedia technologies and the Internet to improve the quality of learning by facilitating access to resources and services as well as remote exchanges and collaboration”.

During the last 10 years a great many institutions worldwide have embarked on developing and offering online distance education. Institutions with a historical background from traditional on-campus education often seem to transfer teaching/learning philosophies, theories, concepts and metaphors from this environment. Keegan argues:

“... that web based education is best regarded as a subset of distance education and that the skills, literature and practical management decisions that have been developed in the form of educational provision known as ‘distance education’ will be applicable mutatis mutandis to web based education. It also follows that the literature of the field of educational research known as distance education, is of value for those embarking on training on the web.”[6]

The above lines point that the skills, research literature, and management solutions developed in the area of distance education are of specific values during the design and developing online e-Learning education systems with better quality. The great emphasis on student support measures developed by leading distance education institutions should be acknowledged when developing the student support systems of future web based e-Learning.

From above definitions the e-Learning may be represented as:

“An interactive and learner centred cloud learning, which includes the delivery of content via Internet, audio/video facilities. The main intention of the e-Learning is to improve the learning performance of learner depending upon the learner’s learning profile”.

3. NEED FOR E-LEARNING

There are many reasons for the need of e-Learning. e-Learning is seen as desirable from a number of perspectives. We outline some of these below.

Shortage of Teacher: In many disciplines, shortage of qualified teachers is a major problem that plaguing most of the educational institutions. This problem will also affect the openings of the new courses. The quality of the available teachers is another major concern. Given the financially attractive opportunities in the industry and poor academic environment that is seen in most of the educational institutions, teaching job is among the lowest in the preference list for many. While hardly anyone looks at e-Learning as an alternative to traditional teaching, in this context, it is seen to expand the reach of the available teachers.

A3 (any time, any place, any pace) learning. For many, the need to come together at a fixed place and time is a major constraint. This is particularly true for those pursuing courses in part-time mode, the just in-time learners, adult learners, etc. The freedom to connect to the course setup at any time of your
choice, and from any place is a major incentive for e-Learning. It also enables learners to take to studying when they feel is the best time for them to study, and hence provides for adapting the TLP to the learner’s individual characteristics.

Enhanced learning experience: This is a very important, but often ignored and under explored aspect. When exploited effectively, e-Learning enables a high degree of personalization and a wide range of instructional methods. Powerful simulation environments, multimedia capability and high-end visualization support enables a learner to relate to the subject much more deeply and hence understands well. Online assessments provide a world of opportunities to further enhancement of the teaching learning process (TLP).

Content creation: India’s contribution of content in the Web is still very poor. One of the main reasons for this is that, very few of our teachers are online. While we happily use online courseware from sources such as MIT Open Course Ware (OCW), we rarely consider contributing our work to share with the world.

Enhancing quality of teaching: When one gets into practicing e-Learning to any significant degree, one will be creating much of the course material electronically. These are a lot more reusable compared to written notes, used otherwise. These can also be shared with other teachers, can be improved over the years using user feedback, and hence results in better quality of material.

More systematic feedback and evaluation: Bringing assessment and other activities under e-Learning, enables us to gather much more detailed feedback on various aspects of the course. These include quality of questions, quality of content, qualitative judgment on students’ performance, etc. These can be used to enhance the quality of instruction at an institutional level.

From the above discussion, one can see a lot of benefits that e-Learning can bring in and learning offers a rich set of options to choose from. Proper choice of these options is important to reap the intended benefits.

4. ASPECTS OF E-LEARNING

The Figure 1 shows the aspects of e-Learning. The main components of e-Learning are the teacher, the learner, the content, the assessment mechanism, communication and collaboration mechanisms, and the administrative aspects.

The teacher and learner

In traditional learning environments, teachers are often the ‘sages on the stage’ the undisputed authority on the subject. For any queries, the learner approaches the teacher. In e-Learning, the students have access to much richer sources of information than the teacher - the internet resources and the vast amount of expertise available thus. e-Learning parlance visualizes the role of a guide by the side. This enables the teacher to reduce the laborious ‘content delivery’ component in teaching, and focus on the needy cases those who are unable to cope with the class and those who are ahead of the class. The availability of detailed information about the students’ progress, the record of the students learning activities, etc online enables the teacher to understand the students on a personal basis in more detail, thus, enhancing the quality of interventions.

Content

The content can be defined as “what is communicated to the learner, directly or indirectly”. The content includes lectures, papers, books, etc. Content is designed to meet some specific objectives and is the most important part of a TLP. It has multiple aspects like creation, management, retrieving, processing, filtering, delivery, etc.

Assessment

When a task is given to a student (i.e. a question, an assessment, etc.) and he responds to it, the quality of response is determined or at least influenced by a number of factors. These include:

- The complexity of the question compared to the competence expected of the learner,
- the quality of teaching imparted relating to the topics relevant to answering that question,
- the level of the student’s understanding of the topics & the clarity of the question- ambiguity, ease of understanding, availability of relevant parameters, etc.

Apart from the final response of a student, one can monitor a number of other parameters which can contribute to computing the feedback. For example, we could record the order in which a student answers the questions in an examination, the number of times he visits a question, the number of times he changes the answer, total time spent on the question and so on. All these provide useful clues to estimating the various aspects mentioned above.

The process of assessments online has five stages:

- Creating and managing a question bank online,
- Creating a question paper from such question bank(s),
- Allowing students to answer this paper online
- Evaluating the answers with computer helpfully automatic, semi-automatic, etc., and
- Post assessment analysis.
One can potentially adopt use of computers for any subset of these stages. For each stage, there are issues, advantages and disadvantages.

Communication and Collaboration
There are different types of communications in a learning environment that can be observed. The primary members of the communication process are the teachers and students. 

Teacher and Student: This is a one-to-one interaction where the student asks a doubt/query and the teacher responds direct to the student. Teacher’s feedback and guidance regarding the students’ performance is also an example.

Teacher and Students: This is one to many or many to one interaction depending on the source of the communication. Common scenarios which include teachers offering clarification on some query to the whole class, announcements about projects/exams, etc.

Student to Student: This is a many to many interaction in general, though sometimes it can be a one-to-one interaction also. For example discussions involving groups of students, debate, etc. Sometimes these may be guided or moderated by a teacher. These types of communications form an integral component of any educational environment, and naturally one needs corresponding mechanisms for these in an e-Learning environment as well.

Discussion board: The most common mode of communication is discussion board (also called bulletin board). Users can post queries, announcements, comments, etc on the board, and they become visible to everyone else. Others can comment on them or seek clarification or offer additional inputs. In general, it is a many to many communication forum. Normally, bulletin boards are threaded linking a chain of messages together, and allow users to navigate the space of messages in different ways.

E-mail: E-mail is often used when there is one to one communication required or when messages are more urgent (since discussion boards are normally checked less often than mails).

Text-chat: Sometimes a synchronous text-chat is also used to support effective discussion sessions. Unlike traditional verbal communication, using electronic media to communicate provide an effective way to record the communications and to share them with other users. This is useful when a teacher is offering clarification to a student, since such information is likely to be of interest to others as well.

One concern in use of these tools is the largely impersonal nature of the communication. The human touch, the non-verbal cues, etc are missing and messages can be misinterpreted. There are netiquette or guidelines which should be understood and practiced to make effective use of these technologies. For example, messages should include emoticons/notations indicating the mood of the sender, such as smiley, expression of surprise, etc., wherever appropriate so that the reader can parse the message in the right spirit.

Administrative aspects
From above discussion an effective deployment of e-Learning would address many aspects including content creation, content delivery, assessment, collaboration and communication between faculty and student(s), and among students etc. In addition, support for online assessment would need mechanisms to keep track of marks per student per assessment, and also computing suitably weighted total. A learning management system, LMS, is an integrated application that provides all these under one umbrella. With one login, one can see and access all relevant aspects of a course. From an administrative perspective, the mechanisms to control access to the course allowing valid students and faculty are needed.

5. LIMITATIONS OF E-LEARNING

The e-Learning has a lot of advantages over traditional learning although, it has limitations. Few of them are,

- Slow access to the course content due to technology limitations,
- Non adaptive content delivery,
- Interactivity is difficult, and require additional technologies at both client and server sides.
- Technological barriers such as internet coverage and limited bandwidth.
- Security of the course content and content delivery, and
- Assessment of learners’ progress is static over different class of learners.

6. CONCLUSION

The e-Learning is a means of cloud learning, that utilizes Internet and audio/video facilities for content delivery to a learner depending upon his learning profile. Recently e-Learning is being implemented by many schools, colleges and Universities to solve various academic problems which include shortage of teachers, poor performance of the students, distribution of study material and administrative problems. With few limitations, e-Learning may be a better solution for better education.

REFERENCES


E-Learning: A Modern Teaching-Learning Environment


[18] Rekkedal, T. 1998: *Courses on the WWW - Student Experiences and Attitudes Towards WWW Courses* - II

CHANNEL MODELING AND PERFORMANCE ANALYSIS OF SHIFT KEYING MODULATION TECHNIQUES IN VARIOUS FADING CHANNELS OF WIRELESS BODY AREA NETWORKS

F.A. DURGA PRASAD¹, S. B. NIRANJAN N CHIPLUNKAR² & T. C. K PRABHAKAR NAYAK³

¹Manipal Institute of Technology, Manipal, Karnataka, India.
²NMAM Institute of Technology, Nitte, Karnataka, India, Visvesvaraya Technological University, Belgaum, Karnataka,
³Electronics and Communication Engineering department, Manipal Institute of Technology, Manipal, Manipal University, Karnataka, India.

Abstract—Wireless Body Area Networks (WBAN), also known as Wireless medical telemetry provides the benefits of the unconstrained freedom of movement for the patients. This system connects individual nodes that are situated in the on body and in body of a patient through a wireless communication channel. The parameters related to human body such as heart rate, temperature, Electro Cardio Gram (ECG), Electro Encephalo Gram (EEG) in terms of Radio Frequency (RF) wave from the body sensor nodes are propagating as on body or in body sensors. Channel models can be used to evaluate the energy efficiency of different network topologies. Hence the propagation mechanism is very complex in comparison with the other wireless channels. Hence channel models for these environments are distinctive. The proposed channel models can be used to evaluate the energy efficiency of different network topologies. This characterizes the physical layer of WBAN which is an important element in the development of a WBAN. In order to achieve higher data rate and low power consumption by the system, shift keying methods of modulations such as Binary phase shift keying (BPSK), Differential phase shift keying (DPSK) and Quadrature phase shift keying (QPSK) etc. are adopted in the physical layer of WBAN. In order to evaluate the performance of the modulation schemes, bit error rate (BER) analysis has been performed considering the wireless channel between various nodes with fading channels such as Rayleigh, Rician and AWGN channels.. The simulation results reveal that Binary phase shift keying (BPSK) scheme shows lower bit error performance in AWGN, Rayleigh and Rician channels than Frequency shift keying (FSK) and Quadrature amplitude modulation (QAM).

Index Terms— Bit error rate, channel model, fading, UWB.

I. INTRODUCTION

In today's world, wireless communication field is growing at a fast rate. Future communication systems are driven by the concept of being connected anywhere at any time. This is not limited to even in medical healthcare field. It is made possible to monitor the health of a patient because of the radio technology and various sensors. Wireless medical communications assisting the variations in the health of people by replacing wires in a hospital are the applications of wireless communications in medical healthcare. Wireless medical telemetry is an example in which a patient's health is monitored remotely. This improves in quality of patient care and the efficiency of hospital administration capabilities. Moreover, this system also serves the purpose of reducing healthcare costs because it permits the remote monitoring of several patients simultaneously.

The development of this technology has lead to a Wireless Body Area Network (WBAN), where tiny smart wireless medical sensors measuring essential parameters of human body such as temperature, blood sugar level sensors, Electro Cardio Gram (ECG), non-invasive blood pressure and the blood oxygen saturation level sensors placed in (implantable) and around (worn close to the body) the body. Communicating devices that are located around human body networked with a coordinator such as Access point are said to be Wireless Body Area Networks (WBAN). These sensors can communicate with the outside world using wireless networks and provide medical information. This information can be forwarded to a physician in real-time [1]. This communication extends to any other distant located devices. Some of the communication technologies available for achieving the above mentioned operation are Bluetooth and Zigbee. Fig 1. Shows the basic block diagram of a WBAN.

A Body Area Network will be a network containing sensor nodes in close proximity to a person’s body (as a cylindrical structure) monitoring vital signals of the human body and a more intelligent node capable of handling more advanced signal processing techniques. It is also called Wearable health monitoring system as the patient wears various sensors that are linked to a common access points. By this convenient means, elderly people can keep track of their health conditions without frequent visits to
their doctors’ clinics. Meanwhile, their doctors can still access the data and give their patients advice based on these data [2]. In order to design and develop a competent and reliable system suitable for WBAN, a knowledge of a radio propagation channel as well as a simple and generic channel model are inevitably required. A new IEEE 802.15 working group for WPAN has developed a channel model in which fading power profile is generated in terms of received power with respect to transmit power, which is based on NICTA’s (National Information and Communication Technology Australia) measurement results at 820 MHz. Hence channel modeling is an essential component of the design of Wireless Body Sensor Networks (WBSN). Since there are different requirements for the systems by various applications, the several technologies for physical layer designs are suggested as on-off keying (OOK), variations of frequency-shift keying (FSK), differential phase shift keying (DPSK), quadrature phase-shift keying (QPSK), Pulse Position Modulation (PPM) and quadrature Amplitude Modulation (QAM) and so on. The ultra wideband (UWB) technologies are also proposed to reduce the power consumption and to achieve high data rate. The behavior of the modulated signal by the above mentioned modulation schemes through different fading channels such as Rayleigh, Rician and AWGN (Additive White Gaussian Noise Channel) are considered. Hence Bit Error Rate (BER) analysis is an essential parameter that measures the behavior of the modulated signal under various channel types [3]. This paper attempts to compare the bit error rate (BER) curves for various modulation schemes through different fading channels. Section II deals with physical layer of the WBAN and Section III discusses the channel modeling and characteristics of the fading channels. The next section (Section IV) explains some of the modulation schemes with the various modulation schemes used in Wireless Body Area networks as well as the measurable properties. In Section V simulation and obtained results are discussed and finally section VI summarizes the paper.

II. PHYSICAL LAYER OF WBAN

Physical layer is the lower most layer of the WBAN system reference model. In the WBAN, the physical layer deals with physical attributes such as the type of signaling, voltage/current levels, transceiver modulation and demodulation methods etc. The two popular energy-efficient technologies of computing models for sensors and other I/O devices employed in a WBAN system that are used to communicate with a more powerful central device are IEEE 802.15.1 (Bluetooth) and IEEE 802.15.4 (ZigBee), in terms of design cost, performance, and energy efficiency. They are the low power radio technology used for achieving wireless transmission of health related parameters. The average power consumption of the radio in the sensor nodes must be reduced below 100µW [1]. But today Bluetooth and Zigbee technologies cannot meet this stringent requirement and new innovative solutions must be found. Hence Ultra-Wideband (UWB) communication is believed to have strong advantages which are promising for WBAN applications. IEEE has launched the Task Group (TG) for wireless body area network (WBAN) in IEEE 802.15 to satisfy the technical trend [4-5]. The purpose of the WBAN is to define new communication standard with physical layer and medium access control (MAC) protocol for both medical and nonmedical applications within 3 m range. According to the diverse applications, data rate should cover from ten kbps up to ten Mbps [4]. The wireless connection between the devices of a WBAN can occur at different frequencies. Often the ISM (Industrial, Scientific and Medical) frequency bands or UWB (Ultra Wideband) are used. The 2.4 GHz band is normally selected for experimentation because it is freely available and most practical existing technology for WBANs works in this band.

Some of the frequency bands used in WBAN systems are 400MHz, 900MHz, and 2.4GHz and UWB frequency range (3.1-5.1 GHz).

III. CHANNEL MODELING OF WIRELESS BODY AREA NETWORK

A channel model estimates the possible decrease in the received signal at the sensor node. It depends the signal propagation via the body area from one node to the other node. The environment surrounding the body is different and therefore, the channel models for WBAN are different from the ones in the other environments. A Fading Channel is known as communications channel which has to face different fading phenomenon, during signal transmission. A fading model decides the quality of the received at the receiver. Bit error rate is a property that features the quality of the signal transmitted through various fading channels. If a fading channel model includes much of the essential parameters which are compensated using suitable techniques, then the received signal will be error free or bit error rate will be close to least value [7]. In wireless systems, the phenomenon of fading is due to multipath propagation of the signal in space. Multipath is the propagation phenomenon that results in radio signals reaching the receiving antenna by two or more paths. Causes of multipath include reflection and refraction, and reflection from terrestrial objects. The effects of multipath include constructive and destructive interference, and phase shifting of the signal. This distortion of signals caused by multipath is known as fading. The following types of fading are normally considered.
A. Channel modeling
Path loss is a phenomenon which may be defined as the difference between transmitted power (dB) and received power (dB) of sensor nodes. Due to the very wide frequency band of a UWB channel (>500 MHz), the path loss is a function of frequency as well as distance between the Tx and Rx. If $PL(f, d)$ denotes path loss as a function of frequency($f$) and distance($d$), then it can be expressed by a product of the terms $PL(f)$ and $PL(d)$ as below. Path loss may be defined as the ratio of transmitted power to the received power

$$PL(d) = \frac{P_t}{Pr}$$

It can be expressed in dB as below,

$$PL(d)\ dB = Pr(dB) - Pr(dB) = - - - - - - - (A )$$

where (A) and (B) are the basic equations for path loss calculations.

$$PL(f, d) = PL(f)PL(d) ------- (1)$$

The frequency dependence of the path loss is given as

$$PL(f) \propto f^{-k} = - - - - - - (2)$$

where $k$ denotes the frequency dependent factor determined by the geometric configuration of the objects. The distance dependence of the path loss in dB is written as below.

Pathloss for a reference distance of $d_0$ from the transmitter, is given by

$$PL(d_0) = \frac{P_t}{Pr(d_0)} = \left(\frac{4\pi^2d_0^2}{Gr Gr} \right)^{-k}$$

Pathloss for a distance of $d$ from the transmitter, is given by

$$PL(d) = \frac{P_t}{Pr(d)} = \left(\frac{4\pi^2d^2}{Gr Gr} \right)^{-k}$$

Dividing (4) by (3) we get,

$$PL(d) = \frac{d^2}{d_0^2}PL(d_0)$$

$$PL(d)[dB] = 10n \log_{10} \frac{d}{d_0} + PL(d_0)[dB]----------- (5)$$

where $n$ = path loss exponent of the surrounding media.

But for a WBAN, the variation of path loss is a random event and hence,

$$PL(d)[dB]=PL(d_0)[dB] + n \log_{10} \left(\frac{d}{d_0}\right)^2 + X_e = - - - - (6)$$

where $X_e$ = Mean log normal distribution,

We have scattering parameter for a 2 port network given by

$$S_{21} = \frac{b_2}{a_1} \left(\frac{Pr}{Pr}\right)$$

$$= \frac{PL(d)}{-S_{21}} = - - - - - - - - (7)$$

Since

$$\frac{1}{S_{21}} = ||S_{21}||$$

But for wireless body area channel the expression for path loss can be,

$$PL(d) = PL_0 + 10n \log_{10} \left(\frac{d}{d_0}\right) + X_e = - ||S_{21}||$$

where $d_0$ is the reference distance, $PL_0$ is the path loss at the reference distance, $n$ is the path loss exponent ($n = 2$ for free space), and $X_e$ is a shadowing (large-scale) fading defined as the variation of the local mean around the path loss. The parameter $S_{21}$ is the scattering coefficient of a 2-port network.

B. Types of Small scale fading
In small scale fading, the signal moves over a distance in the order of wavelength that leads to rapid fluctuation of the phase and amplitude of the signal.

In wireless communication, there are many models that describe the principle of small-scale fading. Out of these models, Rayleigh fading, Ricean fading and AWGN fading models are most commonly used. a). Rayleigh fading model: The Rayleigh fading is primarily caused by multipath reception. Rayleigh fading is a statistical model for the effect of a propagation environment on a radio signal. It is a good model for describing the effect of heavily built-up urban environments on radio signals. This fading model is most appropriate when there is no line of sight between the transmitter and receiver.

b). Ricean fading model: The Ricean fading model is similar to the Rayleigh fading model, except that in Ricean fading, a strong dominant component is present. This dominant component is a stationary (non-fading) signal and is commonly known as the LOS (Line of Sight Component).

c). Additive White Gaussian Noise Model: The simplest radio environment in which a wireless communications system will have to operate is the Additive-White Gaussian Noise (AWGN) environment. Additive white Gaussian noise (AWGN) is the commonly used to transmit signal while signals travel from the channel and there is a need for simulating background noise of channel.

If $r(t)$ is the received signal, the mathematical expression for the received signal that passed through the AWGN channel is $r(t) = s(t) + n(t)$. Here $s(t)$ is the transmitted signal and $n(t)$ is background noise [7].

III. Modulation schemes of Wireless Body Area Networks
In order to communicate a message signal from one point to another point in a communication channel, whose frequency spectrum does not fall within that fixed frequency range, or one that is unsuitable for the channel, is to change a parameter of the transmittable signal according to the information in the message signal. This method of altering the signal to be transmitted is called modulation. At the other point, the receiver then retrieves the original signal through a process called demodulation.

A. Properties of Modulation techniques
The modulation techniques are expected to have the following three properties:

(i) Good Bit Error Rate Performance: In the presence of fading, interference and thermal noise the type of modulation schemes should achieve low bit error rate.

(ii) Power Efficiency: Power limitation is one of the critical design challenges in portable and mobile applications. Nonlinear amplifiers are usually used to increase power efficiency. However, nonlinearity may degrade the bit error rate performance of some modulation schemes. Constant envelope modulation techniques are used to prevent the growth of spectral side lobes during nonlinear amplification.
(iii) Spectral Efficiency: The modulated signals power spectral density should have a narrow main lobe and fast roll-off of side lobes. Spectral efficiency is measured in units of bit /sec/Hz. WBAN system, especially implant system, should have long battery life. Hence, the simple and efficient modulations are proposed which are Pulse amplitude modulations (PAM), Pulse position modulation (PPM), Binary phase shift keying (BPSK), differential phase shift keying (DPSK), and quadrature phase shift keying (QPSK). However Gaussian FSK (GFSK), Gaussian minimum shift keying (GMSK), and phase silence shift keying (PSSK) are the recent modulations are well-known and have relative advantages in bandwidth efficiency, side lobe reduction, and easiness for implementation for their own modulation characteristics. GFSK and GMSK have power and bandwidth efficiency, respectively. Moreover, DPSK has an advantage that the receiver does not require estimation overheads for channel and carrier phase.B. Various methods of modulation The following are some of the modulation schemes used in WBAN.

(i) QAM is the encoding of the information into a carrier wave by variation of the amplitude of both the carrier wave and a quadrature carrier that is 900 out of phase with the main carrier in accordance with two input signals. That is, the amplitude and the phase of the carrier wave are simultaneously changed according to the information you want to transmit. In 16-state Quadrature Amplitude Modulation (16-QAM), there are four I values and four Q values. This results in a total of 16 possible states for the signal. It can transition from any state to any other state at every symbol time. Since 16 = 24, four bits per symbol can be sent. This consists of two bits for I and two bits for Q. The symbol rate is one fourth of the bit rate. So this modulation format produces a more spectrally efficient transmission. It is more efficient than BPSK, QPSK or 8PSK. QPSK is also known as 4-QAM.

(ii) Differential phase shift keying (DPSK): A common form of phase modulation conveys data by changing the phase of carrier wave. In Phase shift keying, High state contains only one cycle but DPSK contains one and half cycle. Differential Shift Keying is a modulation technique that codes information by using the phase difference between two neighboring symbols. In the transmitter, each symbol is modulated relative to the previous symbol and modulating signal, for instance in BPSK 0 represents no change and 1 represents +180 degrees. In the receiver, the current symbol is demodulated using the previous symbol as a reference. The previous symbol serves as an estimate of the channel. A no change condition causes the modulated signal to remain at the same 0 or 1 state of the previous symbol. Differential modulation is theoretically 3 dB poorer than coherent. This is because the differential system has 2 sources of error: a corrupted symbol, and a corrupted reference.

C. Bit Error Rate

The BER, or quality of the digital link, is calculated from the number of bits received in error divided by the number of bits transmitted. Equation (10) provides the expression for Bit error rate.

\[
\text{BER} = \frac{\text{Bits in Error}}{\text{Total bits received}} \quad \text{--- (10)}
\]

In digital transmission, the number of bit errors is the number of received bits of a data stream over a communication channel that has been altered due to noise, interference, distortion or bit synchronization errors. The BER is the number of bit errors divided by the total number of transferred bits during a particular time interval. BER is a unit less performance measure, often expressed as a percentage. Noise affects the BER performance. Quantization errors also reduce BER performance, through incorrect or ambiguous reconstruction of the digital waveform. The accuracy of the analog modulation process and the effects of the filtering on signal and noise bandwidth also effect quantization errors. BER can also be defined in terms of the probability (POE) [7] as given below in equation(11).

\[
\text{POE} = \frac{1}{2} \left(1 - \text{erf} \left( \sqrt{\frac{E_b}{N_0}} \right) \right) \quad \text{--- (11)}
\]

where erf is the error function, \(E_b\) is the energy in one bit and \(N_0\) is the noise power spectral density (noise power in a 1Hz bandwidth). The error function is different for the each of the various modulation methods. The POE is a proportional to \(E_b/N_0\), which is a form of signal-to-noise ratio. The energy per bit, \(E_b\), can be determined by dividing the carrier power by the bit rate. As an energy measure, \(E_b\) has the unit of joules. \(N_0\) is in power that is joules per second, so, \(E_b/N_0\) is a dimensionless term, or is a numerical ratio.

IV. SIMULATION AND RESULTS

Simulation has been performed for various shift keying modulation systems under different fading environments. Bit error rate has been estimated and plotted with respect to \(E_b/N_0\) values. The modulation schemes considered are PSK, DPSK, FSK, QAM. Initially BER characteristics modulations are these modulation techniques are tested with AWGN channel. Then BER performance through the other two widely used fading models by name Rayleigh and Rician channels are performed. The following figures show the experimental graphs and analysis.
Channel modeling and performance analysis of shift Keying modulation techniques in various fading channels of wireless body area networks

From the Fig.2 it is clear that Phase Shift Keying (PSK) modulation scheme provides low bit error rate than DPSK and FSK types. The performance curves of modulation schemes with Rayleigh and Rician channels show that the variation graph of BER is having less slope than with that of AWGN channels.

From the Fig.4 it is clear that, PSK with $k = 2$ (Constant in the PDF of Rician model) provide minimum BER through Rician channel. Higher the $k$ value, lower will be the bit error rates.

V. CONCLUSION

The paper explains the importance of path loss, channel modeling and bit error rate performance of various shift keying modulation schemes through different channels. It also relates the path loss with reflection loss of a two port network with scattering coefficient $S_0$. The vital physiological body parameter can be modulated in different ways such as on-off keying (OOK), pulse amplitude modulation (PAM), and Pulse position modulation (PPM), Binary phase-shift keying (BPSK), Differential phase shift keying (DPSK), and Quadrature phase shift keying (QPSK or QAM). Simulation has been carried out to analyze the performance of various shift keying modulation schemes in terms of Bit error rate versus Energy/bit to Noise power (Eb/No). Results show that among those, BPSK is the best modulation for AWGN channels, Rayleigh and Rician fading channels.

REFERENCES


Channel modeling and performance analysis of shift Keying modulation techniques in various fading channels of wireless body area networks


BIOGRAPHIES OF AUTHORS:

First A. Durga Prasad was born in Udipi in the year 1972. He graduated in electronics discipline from University Vishveswara College of Engineering, Bangalore, Karnataka, India in 1995. He obtained his Master's degree in power electronics from I.I.M.S. College of Engineering, Bangalore, Karnataka, India in 2000. He has worked in electronics and communication department of N.M.A.M. Institute of Technology, Niti, Karnataka, India since 2000. Presently he is a research scholar at Manipal Institute of Technology, Manipal under Manipal University, Manipal, Karnataka, India. His current research interests include the wireless body area networks and previous research interests are in the field of wireless communication and computer networks.

Second B. Niranj N. Chipulkar graduated in electronics and communication discipline from National Institute of Engineering, Mysore, Karnataka, India in 1986. He achieved his Master’s of technology in computer science from Manipal Institute of Technology, Manipal, Karnataka, India in 1992. He then pursued Ph. D. degree from University of Mysore, Mysore, Karnataka, India in the area of CAD for VLSI.

He was the HEAD of the department of computer science and engineering and currently takes over as PRINCIPAL of N.M.A.M. Institute of Technology, Niti, Udipi district, Karnataka, India. He has published “VLSI CAD” (PH learning publications, New Delhi, in 2011) “CAD for VLSI” (NMAMIT Publications, Niti) in 2007. His current research interests include algorithm design for VLSI, CAD, multilayer architecture and program optimization, Adobe networks, E-learning and advanced microprocessors.

Dr. Chipulkar is a member of IEEE, life member of Computer Society of India (CSI) and Indian Society for Technical Education (ISTE) and fellow of Institute of Engineers (IE), India. He has achieved “Excellent Achievement Award” by Center for International Corporation on Computerization (CICC), Govt. of Japan in 2002.

Third C. K. Prabhakar Nayak graduated in electronics and communication discipline from Manipal Institute of Technology, Manipal, Karnataka, India in 1979. He achieved his Master’s degree in Electronic product design from Indian Institute of Science, Bangalore, Karnataka, India in 1989.
A REVIEW ON FACE RECOGNITION ALGORITHM 
BASED ON SOBEL EDGE DETECTION, DISCRETE WAVELET 
TRANSFORM & CURVELET TRANSFORM

ANUJA BHARGAVA¹, & DEHRADUN²

¹²Uttrakhand Technical University

Abstract -The most successful applications of image analysis and understanding is face recognition. There are at least two reasons for the trend the first is the wide range of commercial and law enforcement application and the second is the availability of feasible technologies. The few methods of face recognition which are generally used are Feature based face recognition method, Sobel edge detection, Discrete wavelet transform and Curvelet transform methods. All have their possibilities and features. In this paper we review and compare recently developed face recognition methods.

Keywords: Face Recognition, Sobel edge detection, Discrete wavelet transform and Curvelet transform

I. INTRODUCTION

Biometric attributes are based on behavioural or physiological features. Physiological are related to the shape of the body, examples are eye pattern recognition, hand palm recognition and facial feature recognition. Behavioural are related to the behaviour of a person, examples are signature dynamics, keystroke dynamics and voice. Compared to other biometrics such as fingerprints and iris human faces have the clear advantage of being natural and non intrusive. As a result of face recognition, used in many applications such as credit card verification, criminal identification, seizure control, security, user login face recognition, human computer interaction, face to find a system, AFRS, album also in uncontrolled environmental conditions, such as variable lighting , a variety of expressions, age, cause etc. The face recognition is one of the most successful technologies which is studied from past 25 years. Recently, a large number of face recognition methods with their modifications have been proposed, and lots of them have made state-of-the-art results under the controlled condition. The various methods of face recognition system are neural network based methods, self organising maps, discrete cosine transform, gabor wavelet transform, several edge detection methods, curvelet transform etc. The typical architecture for the proposed face recognition system is shown in Figure 1. It works in two phases: the first one is learning phase (enrolment), and the second one is authentication phase. In both phases, the proposed system has the following five important modules 1) the sensor module acquires the face biometric, 2) the preprocessing module detects the face 3) feature extraction module process the face biometric data in order to obtain a packed yet discriminative representation of the input face biometric, 4) the matching module compares input feature vectors with stored templates and the resulting matching scores are calculated using Euclidean Distance classifier, and 5) finally, the decision module yields an identification or verification decision based on the matching scores.

Fig 1: Flow chart of face recognition system

In this paper we are comparing the various results of different feature extraction methods. The remainder part of this paper is organized as follows. Section II reviews the various face recognition methods. In Section III, presents the results and comparison analysis. Finally, conclusions are drawn in Section IV.
II. FACE RECOGNITION- A REVIEW

A) Face Recognition using Sobel based edge detection  Compared to other edge operator, Sobel has two main advantages: 1. Since the introduction of the average factor, it has some smoothing effect to the random noise of the image. 2. Because it is the differential of two rows or two columns, so the elements of the edge on both sides has been enhanced, so that the edge seems thick and bright. The face recognition processes by sobel edge detection first perform pre-processing. Pre-processing is performed for image smoothing, which is achieved by applying a low pass Gaussian filter. The Gaussian filter used can be represented as, $h(x, y) = \exp(-x^2 + y^2) / 2\pi \sigma^2$ where, the standard deviation ‘σ’ determines the cut-off frequency of the filter. A smooth $h(x, y)$ preserves the original shape in the image and is less likely to distort the image. In our model we use a non-directional edge detector function for calculating edges. This non-directional edge detector function is based on root mean squared of the horizontal and vertical edge detectors, obtained after applying modified wide convolution kernel based on Sobel’s operator. After applying several directional convolution kernels we identified that the combination of a horizontal and vertical convolution kernel is best suitable for facial edge feature extraction. Let $F(X,Y)$ represent the final image obtained after preprocessing in previous stage. Let us represent the non-directional edge detector function used in this method by $G(X,Y)$. $G(X,Y)$ can be given as,

$$G(X,Y) = \left[ I_x (X,Y)^2 + I_y (X,Y)^2 \right]^{1/2}$$

where, $I_x (X,Y)$ and $I_y (X,Y)$ represent the partial derivative of $F(X,Y)$ with respect to $X$ and $Y$. $I_x (X,Y)$ is estimated as,

$$I_x (X,Y) = \frac{2[f(x+1, y+1) - f(x-1, y+1)]}{4T} + \frac{2[f(x+2, y+1) - f(x-2, y+1)]}{4T} + \frac{3[f(x+1, y) - f(x-1, y)]}{6T} + \frac{2[f(x+2, y) - f(x-2, y)]}{8T} + \frac{2[f(x+1, y-1) - f(x, y-1)]}{4T} + \frac{f(x+2, y-1) - f(x-2, y-1)]}{4T}$$

Similarly $I_y (X,Y)$ is also estimated with respect to $Y$. In the above equation, the scaling factor $1/4T$ is omitted, since the computed derivatives are later compared with a threshold. $I_y (X,Y)$ is also represented as,

$$I_y (X,Y) = \left[ I_x (X,Y) \right]^T$$

Then features are extracted by sobel edge detection method results in

Once $G(X,Y)$ is calculated it has to be compared with a suitable threshold value. One of the outcomes of applying gradient based techniques is thick edges. To deal with the issue of thick edges the threshold used for edge detection has to be chosen appropriately.

$$\text{Threshold} = [1.10 * \text{Mean} (G(X,Y))]^{1/2}$$

By observing the images we can find that our method identifies the broken and missing edges more clearly particularly in the chin region. But we can also notice some noise in the image near hair boundary. This is the residual effect of using a wider kernel. This noise sometimes exists even after applying the thinning algorithms. This is one of the drawbacks of using a wider kernel. But with superior noise reduction algorithms in post processing stage we should be able to suppress this noise. Sanqiang Zhao [1] have proposed a sobel edge method, based on Local Binary Pattern(LBP). LBP[2] actually encodes the different features of an image into the micro-level information of edges, spots and other local features. Based on this observation, Sobel operator with LBP can enhance the local features, and thus more detailed information can be extracted from LBP operation. This new operator is named as Sobel-LBP. Because Sobel operator [3] is very easy and efficient to implement, it only slightly increases the computational work of LBP feature extraction process. Sobel-LBP noticeably improves the average recognition rate of LBP by 35.2%. Wenshuo Gao [4] proposes the Sobel edge detection operator based on soft-threshold wavelet de-noising which can remove salt and pepper noise. After adding Gaussian white noises to the image, using the traditional operators to do edge detection to the images we will detect all the noise points and will also blur the edge details of the images. The first derivative’s extreme range of the adjacent pixels edge will change automatically so they can detect image edge. However, when the image adulterates lots of noise signals, there are gray value differences between white noises and image signals, and they can be detected easily. This leads to the poor detection effect of the classical operators. In order to overcome this defect, the paper combines some commonly used de-noising
methods and these classical operators, such as median filter and mean filter de-noising. M. SHARIF propose the combination of sobel edge detection and laplacian to enhance the local features. This combination is quite effective because the edges are insensitive to these variations. Firstly, Sobel operator is applied on original face image as a result of which two images one horizontal and one vertical are produced. Next, apply discrete cosine transform on both original face image and the average face image and computes the mid frequency components of both images by truncating rest of the coefficients. These retained coefficients carry enough information to recognize the face image. After calculating DCT coefficients, the paper concatenates the retained coefficients of both images. Then it supplies these retained coefficients to the Eigen face method to calculate the feature vectors by Eigenface decomposition. This gives result with 96.25% efficiency.

B) FACE RECOGNITION USING DISCRETE WAVELET TRANSFORM

Another feature extraction technique discrete wavelet transform (DWT) allows a signal to be localized in both time and frequency. It operates by convolving a target function with wavelet kernels to obtain wavelet coefficients representing the contributions of wavelets in the function at different scales and orientations [5, 6]. It has also feature of nature pyramidal decomposition of the data, fast, can be directly applied to images for multi-resolution decomposition. The wavelet transform is expressed as decomposition of a signal \( f(x) = L^2(R) \) into a family of functions, which are translation and dilation of a mother wavelet function \( \Psi_{a,b}(t) \).

\[
\Psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi(\frac{t-b}{a})
\]

where \( a \) is the scale parameter and \( b \) is the translation parameter. \( R \) denotes real number. \( L^2(R) \) denotes the vector space of a measurable, square integral, one-dimensional (1-D) function, where \( L \) denotes the number of decomposition levels, the 2D filter coefficients can be expressed as

\[
\begin{align*}
    h_{LL}(m,n) &= h(m)h(n) \\
    h_{LH}(k,l) &= h(k)h(l) \\
    h_{HL}(m,n) &= g(m)h(n) \\
    h_{HH}(k,l) &= g(k)g(l)
\end{align*}
\]

where, \((m,n)\) is size of the input image, the functions \( h(*) \) and \( g(*) \) in equation 2 represent the coefficients of the high pass (H) and the low-pass (L) filter respectively along the row and column (convolution and down sample) separately. Fig.2 shows the wavelet decomposition of an image. After decomposition, four sub bands, Low-Low pass filter (LL), Low-High pass filter (LH), High-Low pass filter (HL), High-High pass filter (HH) are obtained, which represents the average (A), horizontal (H), vertical (V), and diagonal (D) information respectively.

![Decomposition of wavelet transform](image)

After applying wavelet transform on a texture image, a set of sub-images are obtained at different resolution levels. The proposed system consists of two levels of Cohen Daubechies Feauveau (CDF-9/7) transform and only the second level (LL1) image is used for the analysis as that contains most important texture information. The output from DWT is taken as input to Independent Component Analysis (ICA). Naresh Babu [7] proposed an algorithm with independent component analysis which gives results up to 96%, G.K.Kharate proposed the system, WT is chosen to be used in image frequency analysis and image decomposition because by decomposing an image using wavelet transform and the resolutions of the subband images are reduced then the computational complexity will be reduced dramatically by working on a lower resolution image and wavelet decomposition provides local information in both space domain and frequency domain. Wavelet transform can be performed for every scale and translations, resulting in multiples of scale and translation intervals, resulting in discrete wavelet transform (DWT). The implementation of WT is carried out by applying a one-dimensional transform to the rows of the original image data and the columns of the row transformed data respectively. This gives results up to 94.37% Hazım Kemal Ekenel [8] proposed local wavelet analysis-based face recognition can be performed in two different ways. In the first architecture discrete wavelet transform is performed on the entire face image and then the transformed image is partitioned into non-overlapping rectangular blocks. The important regions for classification are selected with a block selection scheme. The features extracted from each selected block are concatenated to construct the overall feature vector that will be used in classification. In the second architecture, as in the JPEG2000 standard [17], the input face image is first partitioned into non-overlapping rectangular blocks. Then, the block selection is performed to determine on which blocks discrete wavelet transform will be.
A review on face recognition Algorithm based on SOBEL Edge Detection, Discrete wavelet Transform & Curvelet Transform

performed. Then the features extracted from each selected block are concatenated to construct the overall feature vector.

C) FACE RECOGNITION USING CURVELET TRANSFORM

The development of Curvelet Transform by Candes and Donoho in 1999 was motivated by the need of image analysis [10]. Curvelets present highly anisotropic behaviour as it has both variable length and width. At fine scale the relationship between width and length can be expressed as width = length²; anisotropy increases with decreasing scale, in keeping with power law. Second generation curvelet transform [9] has two different digital implementations: curvelets via USFFT (Unequally Spaced Fast Fourier Transform) and curvelets via Wrapping. These new discrete curvelet transforms are simpler, faster and less redundant compared to their first generation version. Both the digital implementations use the same digital coronization but differ in the choice of spatial grid. Curvelets via Wrapping has been used for this work as this is the fastest curvelet transform currently available [9]. Curvelet transform has been developed especially to represent objects with curve-punctuated smoothness [9] i.e. objects which display smoothness except for discontinuity along a general curve. In a two dimensional image two adjacent regions can have differing pixel values. Such a gray scale image will have “edges” i.e. discontinuity along a general curve and consequently curvelet transform will capture this edge information. To form an feature set, collect these interesting edge information which in turn increases the discriminatory power of a recognition system. In this, the images are decomposed into its approximate and detailed components using curvelet transform. These sub-images obtained are called curvefaces which reduces the dimensionality of the original image. Tanaya Mandal and Q. M. Jonathan WuPCA, proposed a method of PCA on curvelet domain. Since by standard eigenface technique [11] each image is first converted to a vector by row (or column) concatenation. Then PCA is applied for dimensionality reduction. Though it provides effective approximation, the method suffers from high computational load and poor discriminatory power [12] which results in 96.6%. Yi-Chun Lee [13] proposed a method consists of three steps i.e Gabor filtering, curvelet transformation and 2DPCA analysis which gives the recognition rate 95.5%.

III. RESULTS AND DISCUSSION

In the complete procedure of face recognition the different methods of feature extraction results are compared, the comparison results of these methods are in Table 1.

<table>
<thead>
<tr>
<th>Different feature extraction method</th>
<th>Recognition Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sobel Edge Detection</td>
<td>35.2%</td>
</tr>
<tr>
<td>Discrete Wavelet Transform</td>
<td>96%</td>
</tr>
<tr>
<td>Curvelet Transform</td>
<td>96.6%</td>
</tr>
</tbody>
</table>

REFERENCES

A review on face recognition Algorithm based on SOBEL Edge Detection, Discrete wavelet Transform & Curvelet Transform

Wavelet Transform 96%  94.37% Curvelet Transform 96.6%


ANALYZING OF VARIOUS HUMAN JOINTS DURING MANUAL MATERIAL HANDLING USING ANSYS AND MATLAB

N. SUNDARAM1 & P. SURESH,2

1Department of Computer, Science and Engineering, Karpagam University, Coimbatore, Tamilnadu, India
2Karpagam College of Engineering(Autonomous), Coimbatore, Tamilnadu, India

Abstract—Painful joints exhibit abnormal motion and vice versa during movements. Most of the patients were suffering from joint pains. This joint pains like Hip joints, Knee joints, Foot joints, Shoulder joints, Elbow joints, and Wrist joints. Patients suffering from joint disorders visit a therapist. The therapist must correlate all these information sources regarding joint problems. Most probable one third of all jobs in industry involve Manual Material Handling (MMH). This Manual Material Handling of human poses risk to many and cause back pain, joint pains and other problems like Knee joints, wrist joints, Shoulder joints, etc. A finite element model is used to study about the stress of human joints. By using Image processing techniques using soft computing like MATLAB and ANSYS. A Biomedical model has been used for optimizing the lifting posture for minimum efforts. This model is also used to predict the lifting material in every individual human being. This study can be extended for loading of muscles.

Keywords—Magnetic Resonance Imaging, Human Joints, Finite Element Modeling, ANSYS, Image Processing.

1. INTRODUCTION

Manual Object handling by humans in industry is one of the important & complex problems like joint pains in the human body. The joints should support the whole weight of the human body. More stress is applied during lifting the object manually. Manual object handling cause cumulative disorders to muscular skeletal system through continuous lifting / handling activities. This work related back pain, human joints, & injuries are the most common muscular skeletal disorders. The human body problems like Hip joints, Knee joints, Shoulder joints, Elbow joints, & Wrist joints affect because of Manual Object handling. A biomechanical model has been used for this problem.

2. LITERATURE SURVEY

Martin Wawro and Madjid Fathi-Torbaghan (2004) observed shows an object-oriented framework for the finite-element (FE)-based simulation of the human knee joint motion. The FE model of the knee joint is acquired from the patients in vivo by using magnetic resonance imaging. The MRI images are converted into a three-dimensional model and finally an all-hexahedral mesh for the FE analysis is generated. The simulation environment uses nonlinear finite-element analysis (FEA) and is capable of handling contact of the model to handle the complex rolling/sliding motion of the knee joint a Parallel Framework for the FE-Based Simulation of Knee Joint Motion.

Ehsan Arabi, Ronan Boulic, (2007) observed a Fast Method for Finding Range of Motion in the Human Joints Finding the range of motion for the human joints is a popular method for diagnosing joint diseases. By current technology, it is more trustable and easier to find the range of motion by employing computer-based models of the human tissues. Range of motion for human joint joints without using any collision its detection algorithm. This method is based on mesh classifying in a cylindrically segmented space. The method shows to be much faster than the traditional ones and provides the accurate results.

Keisuke Shima and Toshio Tsuji (2010) observed the Classification of Combined Motions in Human Joints through Learning of Individual Motions Based on Muscle Synergy Theory. The method can be adopted to represent non-trained combined motions (e.g., wrist flexion during hand grasping) using a recurrent neural network by combining synergy patterns of EMG signals preprocessed by the network. This approach allows combined motions (i.e., unlearned motions) to be classified through learning of individual motions (such as hand grasping and wrist flexion) only, meaning that the number of motions can be increased without increasing the number of learning samples or the learning time needed to control devices such as prosthetic hands. Synergies from EMG signals. As an example, Bizzi et al. tried to extract multiple muscle synergies from time-series EMGs of a frog’s leg, and reported that the measured EMGs could be reconstructed using a combination of the muscle synergies extracted however, the synergies identified focused only on measured EMGs, so cannot be used to estimate unknown motions.

M.M. Mirbagheri1, and K. Settle2 Argentina, (2010) observed that the neuromuscular Properties of Different Spastic Human Joints Vary Systematically Buenos Aires, the mechanical abnormalities of the spastic wrist in chronic stroke survivors, and determined whether these findings were representative of those recorded at the elbow and ankle joints. System identification techniques were
used to characterize the mechanical abnormalities of these joints and to identify the contribution of intrinsic and reflex stiffness to these abnormalities. Modulation of intrinsic and reflex stiffness with the joint angle was studied by applying PRBS perturbations to the joints at different joint angles over the range of motion.

Nareen Karnati, Ben Kent and Erik D. Engeberg (2011) observed Human Finger Joint Synergies for a Constrained Task Applied to a Dexterous Anthropomorphic Hand. Results showed that the periodic motions exhibited by the finger joints shared a common frequency for each subject, but differed in amplitude and phase. From the gathered data, a set of sinusoidal trajectories were developed to approximate this motion for a robotic hand. Because the joint trajectories share the same frequency, a single sinusoidal input can be used in the path planning of the robot to achieve this task. A reference joint is given a sinusoidal input, and the remaining joints are scaled in phase and amplitude with respect to this reference joint. This significantly reduces the computational cost and complexity of the task. Simulation results show that the developed sinusoidal trajectories show a close correlation with the motion profiles seen from human experiments. Constrained motion task are explored.

Nicola Sancisi and Vincenzo Parenti-Castelli (2011) studied about Strip-Driven Devices for the Spatial Motion Guidance of Human Joints Orthoses and exoskeletons need devices that can replicate the natural spatial motion of human joints. These devices should be simple and should have a high accuracy, in order not to constrain and load the joints unnaturally. In this study, strip-driven devices are proposed to guide the spatial joint motion. Classic planar devices are generalized to obtain rolling without slipping between two ruled surfaces. The special case of spherical motion is presented and analyzed in details. The influence of several design parameters on the kinematics and static behavior of these devices is also presented.

3. SEQUENCE OF WORK

Sequence of work can be followed by different stages like CT scan Image, Original Image properties of Input data, parameters, CAD Modeling of 2D and 3D Modeling, Stress analysis using MAT Lab and ANSYS, Image processing, Image Segmentations, Comparison of Results, Extension of work towards Muscles.

4. RESEARCH METHODOLOGY

The Original CT Scanned Image is observed from the medical Lab and taken as input data with their parameters. The input data is converted to 2D and 3D modeling by using CAD techniques. Stress analysis using MAT Lab and ANSYS is applied to the output data. Image processing techniques finds the exact image of both stressed and unstressed image between original CT Scanned image with 2D and 3D image. The images of different segments were taken as results. The results will be compared with original image so that the human joints problems can be analyzed in the human being. Human joints like wrist, elbow, shoulder, hip, knee, and foot joints can be analyzed with this methodology.

5. EXPECTED OUTCOME

Finite element model can predict the loading behaviour of Human joints for the proposed development. The finite element model for the Human joints is to be found. Image processing of Human joints also to be done. The effort to be taken for the in vivo and in vitro data collection and analysis are reduced multi fold in the finite element modelling. By implementing this methodology the worker compensation will be reduced in the industry and loss of productive man-hours will also be decreased. Approximately one third of all jobs in industry involve MMH The study can be extended to include the loading of the muscles. The joints of the lower extremity participate in locomotion. Pathologic conditions that affect the knee impair normal daily activities and the ability to walk. The broad spectrum of pathologic conditions affecting the knee includes...
soft tissue injuries to the ligament us structures that provide stability, meniscal
Injuries and arthritic conditions. The number of patients with pain and disability from arthritic
conditions of the knees is increasing, as are the expectations of those
Who wish to maintain active lifestyles Approximately 300,000 knee replacements are performed each year
in the United States. The development of criminal investigation expert system through practical
elements model is used to study & analyses the success of all human joints using MAT Lab and
ANSYS. For the future work the study is extended to include the loading of the muscles.

6. CONCLUSIONS AND FUTURE WORK

The study shows the problems among human being by having joints paints using manual material
handling. Lifting is considered to be a major cause for low back pain, joints paints and spiral injuries
finite elements model is used to study & analyses the success of all human joints using MAT Lab and
ANSYS. For the future work the study is extended to include the loading of the muscles.

REFERENCES

[1]. Martin Wawro and Madjid Fathi-Torbaghan* A Parallel
Framework for the FE-Based Simulation of Knee Joint Motion VOL. 51, NO. 8, AUGUST 2004
[2]. Gideon F. Inbar Modeling the Human Elbow Joint Dynamics: Estimation of Joint Stiffness with Different Loads and
Movement Velocities 14th 8MESl - 1995
[3]. Wen Cheng Chen, Noah Chiu, J.Y. Lu, Jonathan Yeh, Jessica
Kuan Manual Material Handling System 9, Creation Rd. 1, Hsinchu Science Park Hsinchu, Taiwan 300 [4] Ahhat Wang,
Mingcong Deng Human Arm-Like Robot Control Using the Viscoelasticity of Human Multi-Joint Arm 2-24-16 Nakacho,
Kagonei-shi, Tokyo 184-8588, Japan
[5]. Keisuke Shima and Toshio Tsujii Classification of Combined
Motions in Human Joints through Learning of Individual
Motions Based on Muscle Synergy Theory 978-1-4244-9315-
9/10 ©2010
[6]. M.M. Mirbagheri1,2 and K. Settles2 Neuromuscular Properties
of Different Spastic Human Joints Vary Systematically
Buenos Aires, Argentina, August 31 - September 4, 2010
[7]. Nareen Karnati, Ben Kent and Erik D. Engeberg Human
Finger Joint Synergies for a Constrained Task Applied to a
Dexterous Anthropomorphic Hand Bled, Slovenia, October
26-28, 2011
Local Polynomial Modeling and its application to Image
Smoothing and Reconstruction # The Author(s) 2010.30 May
2010
& Z. G. Zhang & Zhongfu Ye & H. Y. Shum
[10]. The Design and Construction of a Movable Image-Based
Rendering System and Its Application to ultiview
Conferencing 16 November 2010
[11]. Lemart Yseboodt · Michael De Nil · Jos Huiskens · Mladen
Berekovic · Qin Zhao Frank Bouwens Design of 100 I W
Wireless Sensor Nodes for Biomedical Monitoring 17
October 2008.
[12]. L. Dovat, O. Lambercy, R. Gassert, T. Maeder, T. Milner, T.
C. Leong, and E. Burdet, “HandCARE: A cable-actuated
rehabilitation system to train hand function after stroke,”
Cramer, “A robotic device for hand motor therapy after
stroke,” in Proc. IEEE Int. Conf. Rehabil. Robot. (ICORR),
multiple user virtual environments for rehabilitation (MUVER)
algorithm of time varying multi joint human arm
viscoelasticity, Proc. of the International Joint Conference on
Biomedical Engineering Systems and Technologies,
Valencia, Spain, 2010.
[16]. A. Ottoboni, V. Parenti-Castelli, N. Sancisi, C. Belvedere, and
A. Leardini, “Articular surface approximation in equivalent
spatial parallel mechanism models of the human knee joint,”
[17]. Demid Borodin · B. H. H. (Ben) Juurlink · Vassiladiis
DESIGN AND STRESS ANALYSIS OF BIOMECHANICAL 2D MODEL OF FUNCTIONAL LUMBAR SPINE USING A-CAD

S.SENTHILRAJA¹ & P.SURESH²

¹Dept. of Computer Science & Engineering, Karpagam University, Coimbatore, Tamilnadu, India
²Karpagam College of Engineering (Autonomous), Coimbatore – 641 032, Tamilnadu, India.

Abstract - Ergonomic approach to the Manual Material Handling (MMH) tasks like lifting, pushing, pulling is of interest to the industrial ergonomists. MMH especially, lifting, poses a risk to many and considered as prime cause of back-pains and various other joint impairments. Industrial ergonomists place topmost emphasis on optimisation of posture for safe lifting. In vitro and in vivo studies on real subjects are risky and studies on cadaveric FSU are time consuming. In this paper, we have introduced the Mechanical software A-CAD for design the two dimensional model of Lumbar spine for analysis and predict the stress value during lifting of loads.

Key words- Lumbar spine, Biomechanical model, A-CAD, Manual Material Handling

I. INTRODUCTION

Biomechanics provides estimation of various mechanical stresses acting on the body while a person manually handles an object. Estimation of stresses imposed on the body musculoskeletal system based on biomechanical model was found to be successful (Maurel et al, 1996). National Institute For Occupational Safety And Health (NIOSH) recognised the problem of work related back injuries and published the Work Practices Guide (WPG) for manual lifting. Actual biomechanical analysis of lifting exertions often require measured values of applied trunk moments and forces as base line for validation data. Accurate measures of the trunk kinetic data are difficult to achieve from dynamic exertions without significant approximation, cost, or motion constraints. It calls for experimental set up by Granata (1996).

II. LITERATURE REVIEW

Many other universities and institutions around the world are producing finite element models of the spine for running various analyses. For example, there are FEMs being created to study the effects of automatic accidents on the spine. The geometry, material properties and loading conditions of the lumbar spine are very complex. Though the FEM is well-established method, various simplifications have to be used and assumptions are to be made (Bowling et al, 1995). The spine is established by muscle forces. In vitro experiments almost never, consider the muscle forces. It is reported that, experimentation has shown that in lordotic spine during prolonged standing, the impacted joints at each segment level bear an average of 16% of the axial load (Oliver, 1991). Hutchinson and Littlefield (Cantau, 1991) tried to simplify FEM of vertebral body by modelling it as a cylinder. This study was conducted to determine the stresses induced in a previously injured spine during pilot ejection. Initially, Hutchinson and Littlefield started FEM by importing data from the DIGIBOT images of the vertebrae into ALGOR V. After an internal mesh was generated by ALGOR V, a PERL script was written to translate the data for ABAQUS to read. The model was too complicated to be processed. As a result of which they have constructed the simplified model. Shirazi-Adl’s study (1986) was referenced for the material properties of the lumbar region. The empirical quantities of the spinal components used in shirazi-Adl’s study are shown in Table 3.1. Work on FEM analysis of segmental level lumbar L2/L3 and cervical spine (Carlos, 2000: Lian, 2000) were found during the survey. But, the combined L1/L2/L3/L4/L5/S1 (FSU) FEM analyses have not been found to have been done.

III. DESIGN AND ANALYSIS OF FUNCTIONAL SPINAL UNIT (FSU)

Manual Material Handling (MMH) activities, lifting is considered to be a major cause for low back pain and spinal injuries. Annual costs associated with back pain in US alone ranged from 20 billion US$ to 50...
Design and Stress Analysis of Biomechanical 2d Model of Functional Lumbar Spine using A-Cad

billion US$ (Miller, 1997). Although epidemiological studies have suggested possible causes, the actual mechanisms by which the lumbar spine is injured during load cycle resulting in low back pain, remains unknown. However, the response of the disc to loading conditions that occur during lifting is difficult to measure in vivo and in vitro and has not been investigated using any model (Williams, 2002). In vitro and in vivo studies on real subjects are risky and studies on cadaveric FSU are time consuming. Here, an attempt has been made to analyze FSU using suitable finite element modeling for sagittal plane lifting activity. Spinal unit, otherwise known as vertebral unit consists of 24 separate bony vertebrae together with 5 fused vertebrae, which form the sacrum, and usually 4 fused vertebrae, which form the coccyx, 24 separate vertebrae are interlaced with inter-vertebral discs. Normally the column appears symmetrical in the frontal plane and has characteristic curvatures in the sagittal plane. Curvatures provide natural shock absorbency and flexibility (Fig 1). Three principal functions of vertebral column are to
1. Support the human in the upright posture
2. Allow moment and locomotion
3. Protect the spinal cord.

IV. TWO DIMENSIONAL MODELLING

2D modeling is the modeling of a continuous system, which has an infinite number of degrees of freedom, using a representative geometry of that system made up of a finite number of smaller elements and node points. The more elements in the model, the more accurate it is. The material properties, displacements, and other system characteristics are represented by mathematical functions between nodes. This finite element model can then be used to determine the stress, strain, and displacement of the structure resulting from external loading (Cantau, 1997). Once the Finite Element Model has been created and the system characteristics have been established in the model, a global stiffness matrix can then be formed for the whole structure. Given the forces and boundary conditions, the unknown displacements at each of the node points can then be used to determine the stresses and strains acting on each element (Cantau et al., 1997). Initially, the Finite Element Models were applied to aircraft structures, and then FEM rapidly spread to Civil Engineering and Mechanical Engineering. FEM is gaining acceptance as a valuable tool to study static, dynamic as well as cyclic problems. Only recently FEM has been seriously applied biomechanical problems.

ANTHROPOMETRY OF FSU

Functional Spinal Unit (FSU) consists of Lumbar and sacrum region. Lumbar vertebral column comprises of five vertebrae and the intervertebral discs. It has a characteristic curvature called lumbar lordosis. In the standing position, the sacrum is tilted forwards so that its upper surface is inclined forwards and downwards forming an angle between the top of the sacrum and the horizontal, which varies between 50° – 53°

FACTORS CONTRIBUTING TO THE NORMAL SHAPE OF THE LUMBAR LORDOSIS

The common factors that maintain the normal shape of the lumbar lordosis are listed below:
1. L5 vertebral body is wedge shaped, the anterior body wall is 3 mm higher than the posterior body wall. This brings the upper surface of the L5 body closer to the horizontal plane than the upper surface of the sacrum
2. In addition, the L5/S1 disc is also wedge shaped, the anterior vertical height is 6~7 mm greater than its posterior height. As a result of the wedge shape of the disc, the lower surface of the vertebral body is not parallel to the upper surface of the sacrum, so that the angle formed between the two surfaces may vary between 6~29° and has an average size of 16°.
3. Each vertebra above is inclined slightly backwards in relation to the vertebrae below.
4. In 75% of the adults the center of gravity lies anterior to the vertebral column. In these individuals there is constant activity in the erector spine muscles, which work to prevent the trunk from falling forwards and assist in maintaining the lumbar lordosis.
5. In a normal spine, in the upright posture, the body of L5 lies directly vertically above the sacrum. Various attempts have been made to measure the lordosis, but as investigators have used different parameters the result differ substantially.
ASSUMPTIONS

Following assumptions were made while modelling FSU:
1. Vertebrae were considered as having elliptical cross section. Actually this is an improvement over Hutchinson-Littlefield model. Ellipse gives a closer approximation.
2. Upright standing posture is assumed for modelling.

V. MODELLING AND ANALYSIS OF FSU

With the advent of Magnetic resonance imaging (MRI), one can measure internal organs and bones with high resolution and accuracy. MRI is a non-invasive imaging approach. Cross sectional size of L1 lumbar vertebra were measured from an adult subject using MRI scan. It is found to be 40mm Major and 32 mm minor. Based on the anthropometric proportions, other dimensions were arrived at. There are host of software for 3D modelling and here Pro/Engineer is preferred because of its simplicity and seamless integration with analysis package Pro/Mechnica. Whole FSU were modelled in parts and assembled according to the configuration as shown in Fig 2. Assembled model was transferred to Pro/Mechnica environment.

BOUNDARY CONDITIONS

Lower vertebral disc (L5/S1) of the L5 body was fixed. All other vertebral bodies and discs were allowed all degrees of freedom. Loading according to literature, lumbar vertebrae L1 share only 11% of axial load, leaving the rest of the load of muscles. This was taken in to consideration while loading the FE model. Axial force due to body weight is calculated (at L5/S1) as following:

From forces along spine direction

\[ F - ES - B \cos \theta - W \cos \theta = 0 \]

\[ \text{Moment:} \]

\[ ES \times l = Bb + Ww \]

From forces along spine direction (For lifting 10 Kg load)

\[ ES \times 0.06 = (400 \times 0.18) + (98 \times 0.35) = 72 + 34.3 \\
ES = \frac{72 + 34.3}{0.06} = 1771.5 = 1772 \]

Substituting the ES value in equation 1.

\[ F - 1772 - 400 \cos 52^\circ - 98 \cos 52^\circ = 0 \]

F = 2079 N

Where

\[ \theta = 52^\circ \]

Axial force on L1 is 228 N @ 11 %

Load was applied on L1 top surface. Refer Fig 2 Pro/Mechnica’s AutoGEM created elements and meshes. Solver was run to get the results with 2% p-loop convergence. Fig 2 shows the deformation of the model for lifting a load of 10 Kg. Calculations are as shown below: Maximum vonmises stresses are found at L5/S1 and L4/L5 and its magnitude found to be 3.05 Mpa.

VI. CONCLUSION

The model assembly of the FSU of this study was done using PRO/E. FSU assembly model is shown in Fig 2. As it was mentioned, the model was first generated for upright position with only body weight condition. The normal posture of the FSU was emulated to the reality. The segments from L1 through L5 were modelled for a normal adult. Fig 24.5 shows the Finite Element Model of the FSU under loaded condition and the constraints. The deformation of the model, under the load of 10 Kg and boundary condition, is depicted. Maximum Von Mises stress was found at L5/S1 and at L4/L5 and its magnitude was found to be 3.063 MPa.

REFERENCES


Design and Stress Analysis of Biomechanical 2d Model of functional Lumbar Spine using A-Cad


[6]. Jenna Bowling, Tony Chao and Robin Kinsey (1995), 'Analysing the spine under aircraft ejection loading', Project reports. The University of Texas at Austin.


A LAYERED DECODING ARCHITECTURE FOR LDPC DECODER WITH LOW ENERGY CONSUMPTION

G.SURESH KRISHNA SAGAR¹ & M.SANDEEP KUMAR,²
¹GATES Institute of Technology, Anantapur
²GITAM university, Vizag

Abstract—Low-density parity-check (LDPC) decoder requires large amount of memory access which leads to high energy consumption. To reduce the energy consumption of the LDPC decoder, memory-bypassing scheme has been proposed for the layered decoding architecture which reduces the amount of access to the memory storing the soft posterior reliability values. In this work, we present a scheme that achieves the optimal reduction of memory access for the memory bypassing scheme. The amount of achievable memory bypassing depends on the decoding order of the layers. We formulate the problem of finding the optimal decoding order and propose algorithm to obtain the optimal solution. We also present the corresponding architecture which combines some of memory components and results in reduction of memory area. The proposed decoder was implemented in TSMC 0.18 m CMOS process. Experimental results show that for a LDPC decoder targeting IEEE 802.11n specification, the amount of memory access values can be reduced by 12.9–19.3% compared with the state-of-the-art design. At the same time, 95.6%–100% hardware utilization rate is achieved.

Index Terms—Low power, low-density parity-check code, simulated annealing.

I. INTRODUCTION

Recently, low density parity check codes (LDPC) have gained significant attention due to near Shannon limit performance. They have been adopted in several wireless standards such as DVB S-2, IEEE 802.16e and 802.11n because of their excellent error correcting performance. A LDPC code is a linear block code defined by a sparse parity check matrix [Fig. 1(a)]. It can be represented by a bipartite graph, called Tanner Graph as shown in Fig. 1(b), which contains two sets of nodes: variable nodes that represent the bits of a codeword and check nodes that implement the parity-check constraints. The standard decoding procedure is the message passing algorithm, also known as “sum-product” or “belief propagation” (BP) algorithm, which iteratively exchanges the messages between the check nodes and the variable nodes along the edges of the graph. In the original message passing algorithm, the messages first are broadcasted to all the check nodes from the variable nodes and then along the edges of the graph the updated messages are fed back from the check nodes to the variable nodes to finish one iteration of decoding. There

![Fig. 1. Example of parity check matrix of a LDPC code and its Tanner graph representation.](image)

are different ways to implement the LDPC decoder, based on the number of processing units available. In general, the LDPC decoder architecture can be classified into three types: fully parallel architecture, serial architecture and partial parallel architecture. In fully parallel architecture, a check node processor is needed for every check node, which usually results in large hardware cost and complicated routing, and hence is less flexible. The serial architecture uses just one check node processor to share the computation of all the check nodes and is too slow for most applications. For partial parallel architectures, multiple processing units are used allowing proper tradeoff between the hardware cost and the throughput and are commonly adopted in the actual implementation. In order to achieve higher convergence speed, i.e., to minimize the number of decoding iterations, serial message passing algorithm, also known as layered decoding algorithm, has been proposed together with the corresponding partial parallel architecture. There are two types of layered decoding schemes: vertical layered decoding and horizontal layered decoding. In the horizontal layered decoding, a single or a certain number of check nodes (called layer) are first updated. Then the whole set of neighboring variable nodes are updated, and the decoding process proceeds layer after layer. Dually, in the vertical layered decoding, a single or a certain number of variable nodes (layer of variable nodes) may be updated first. Then the whole set of neighboring check nodes are updated. Because the serial check node processor is easier to be implemented in VLSI and therefore the horizontal layered decoding is preferable for practical implementations because of the faster convergence and regular architecture, layered decoders are commonly found in the LDPC decoder implementation. In this work, we focus on the LDPC decoding implementation based on the layered decoding algorithm. High power consumption is one of the bottlenecks for LDPC decoder design. Recently low power design techniques and
architectures have been proposed for fully parallel, and partial-parallel LDPC decoder architecture. The partial parallel architectures based on the layered decoding algorithm have efficiently reduced the hardware cost and sped up the convergence rate, and already led to energy-efficient design comparing with other architectures. However how to further reduce its energy consumption is still a challenging design problem. Due to the large amount of memory access, the power consumption of the memory access accounts for the major part of the total power consumption of the layer decoder. Reducing the energy consumption of the memories is the key issue to realize a low energy LDPC decoder. At the algorithmic level, the Min-sum decoding algorithm and its variants have been proposed, which greatly reduces the memory storage required for the check to variable messages, and also the energy consumption of the memories of the LDPC decoder with insignificant performance loss. At the architectural level, memory-bypassing scheme has been proposed in to reduce the amount of the memory access by utilizing the characteristic of the LDPC parity check matrix and the decoding algorithm for the layered decoding architecture. However, the scheme proposed in may not lead to optimal reduction of memory access. In this work, we propose several schemes that lead to the optimal reduction of the memory access. Compared with the previous works, this work has the following contributions. Firstly by de-coupling the read and write access order of the memory storing the soft posterior reliability values (we denote it as the Channel RAM), optimal amount of memory bypassing is achieved and the number of idle clock cycles is reduced. Secondly, we are not only considering memory bypassing between consecutive layers of decoding but extending it to multiple layers. By doing so, the number of memory bypass is maximized. Thirdly, the decoding order of the layers has significant impact on the amount of memory bypass that can be achieved. We formulate the problem of finding the optimal decoding order given a parity check matrix as a searching problem and propose algorithm to obtain the optimal decoding order. Fourthly, based on the decoupling of the read and write order of the Channel RAM, we propose a memory efficient architecture, in which the Channel RAM and the memory storing the intermediate data are merged into a single memory. By doing so, the overall area is reduced. The rest of the paper is organized as follows. Section I gives the background of LDPC decoding scheme, the traditional layered LDPC decoder architecture and the memory-bypassing scheme for the layered decoding architecture proposed in [10]. The proposed memory bypassing scheme is presented in Section III. A quick searching algorithm to find the optimal decoding order of the layers of the LDPC base matrix, which results in the maximum overlapping, is described in Section IV. The overall decoder architecture implementing the memory-by-passing scheme is described in Section V. In Section VI, experimental results and comparisons among different LDPC decoders are presented. Conclusions are drawn in Section VII, comparing with other architectures. However how to further reduce its energy consumption is still a challenging design problem. Due to the large amount of memory access, the power consumption of the memory access accounts for the major part of the total power consumption of the layer decoder. Reducing the energy consumption of the memories is the key issue to realize a low energy LDPC decoder. At the algorithmic level, the Min-sum decoding algorithm and its variants have been proposed, which greatly reduces the memory storage required for the check to variable messages, and also the energy consumption of the memories of the LDPC decoder with insignificant performance loss. At the architectural level, memory-bypassing scheme has been proposed in to reduce the amount of the memory access by utilizing the characteristic of the LDPC parity check matrix and the decoding algorithm for the layered decoding architecture. However, the scheme proposed in may not lead to optimal reduction of memory access. In this work, we propose several schemes that lead to the optimal reduction of the memory access. Compared with the previous works, this work has the following contributions. Firstly by de-coupling the read and write access order of the memory storing the soft posterior reliability values (we denote it as the Channel RAM), optimal amount of memory bypassing is achieved and the number of idle clock cycles is reduced. Secondly, we are not only considering memory bypassing between consecutive layers of decoding but extending it to multiple layers. By doing so, the number of memory bypass is maximized. Thirdly, the decoding order of the layers as significant impact on the amount of memory bypass that can be achieved. We formulate the problem of finding the optimal decoding order given a parity check matrix as a searching problem and propose algorithm to obtain the optimal decoding order. Fourthly, based on the decoupling of the read and write order of the Channel RAM, we propose a memory efficient architecture, in which the Channel RAM and the memory storing the intermediate data are merged into a single memory. By doing so, the overall area is reduced. The rest of the paper is organized as follows. Section I gives the background of LDPC decoding scheme, the traditional layered LDPC decoder architecture and the memory-bypassing scheme for the layered decoding architecture proposed in [10]. The proposed memory bypassing scheme is presented in Section III. A quick searching algorithm to find the optimal decoding order of the layers of the LDPC base matrix, which results in the maximum overlapping, is described in Section IV. The overall decoder architecture
implementing the memory-by-passing scheme is described in Section V. In Section VI, experimental results and comparisons among different LDPC decoders are presented. Conclusions are drawn in Section VII, comparing with other architectures. However, how to further reduce its energy consumption is still a challenging design problem. Due to the large amount of memory access, the power consumption of the memory access accounts for the major part of the total power consumption of the layer decoder. Reducing the energy consumption of the memories is the key issue to realize a low energy LDPC decoder. At the algorithmic level, the Min-sum decoding algorithm and its variants have been proposed, which greatly reduces the memory storage required for the check to variable messages, and also the energy consumption of the memories of the LDPC decoder with insignificant performance loss. At the architectural level, memory-bypassing scheme as been proposed in to reduce the amount of the memory access by utilizing the characteristic of the LDPC parity check matrix and the decoding algorithm for the layered decoding architecture. However, the scheme proposed in may not lead to optimal reduction of memory access. In this work, we propose several schemes that lead to the optimal reduction of the memory access. Compared with the previous works, this work has the following contributions. Firstly, by de-coupling the read and write access order of the memory storing the soft posterior reliability values (we denote it as the Channel RAM), optimal amount of memory bypassing is achieved and the number of idle clock cycles is reduced. Secondly, we are not only considering memory bypassing between consecutive layers of decoding but extending it to multiple layers. By doing so, the number of memory bypass is maximized. Thirdly, the decoding order of the layers as significant impact on the amount of memory bypass that can be achieved. We formulate the problem of finding the optimal decoding order given a parity check matrix as a searching problem and propose algorithm to obtain the optimal decoding order. Fourthly, based on the de-coupling of the read and write order of the Channel RAM, we propose a memory efficient architecture, in which the Channel RAM and the memory storing the intermediate data are merged into a single memory. By doing so, the overall area is reduced. The rest of the paper is organized as follows. Section I gives the background of LDPC decoding scheme, the traditional layered LDPC decoder architecture and the memory-bypassing scheme for the layered decoding architecture proposed in [10]. The proposed memory bypassing scheme is presented in Section III. A quick searching algorithm to find the optimal decoding order of the layers of the LDPC base matrix, which results in the maximum overlapping, is described in Section IV. The overall decoder architecture implementing the memory-by-passing scheme is described in Section V. In Section VI, experimental results and comparisons among different LDPC decoders are presented. Conclusions are drawn in Section VII.

\[ \mathbf{H} \cdot \mathbf{x}^T = 0 \quad \forall \mathbf{x} \in C. \]  

(1)

The node can also be described as means of a bipartite graph, known as Tanner graph (Fig. 1(b)) [6]. A Tanner graph is made up of two entities, variable nodes (VN) and check nodes (CN), and they are connected to each other through a set of edges. An edge links the check node \( m \) to the variable node \( n \) if the element \( H_{m,n} \) of the parity check matrix \( \mathbf{H} \) is non-null. The optimal LDPC decoding is achieved by using a message passing algorithm, also known as "belief propagation" (BP), which can be described as an iterative exchange of messages along the edges of the Tanner graph. The algorithm proceeds iteratively until a maximum number of iterations are elapsed or a stopping rule is met. Inputs of the algorithm are the intrinsic Log-Likelihood Ratios (LLRs) of the received bits (i.e., the variable nodes) also referred to as a priori information.

The following describes the belief propagation algorithm: 

\[ Q_{m,n}(q) \] is the check-to-variable message from check node \( m \) to variable node \( n \) in the \( q \)th iteration. \( \lambda_{m,n} \) is the variable-to-check message from variable node \( n \) to check node \( m \) in the \( q \)th iteration. \( \Lambda_{m}^{[q]} \) is the set of the neighboring variable nodes of variable node \( n \) and \( N_m \) is the set of the neighboring variable nodes of check node \( m \). In the \( q \)th iteration, the variable node process and the check node process are computed as follows.

- **Variable node process**: the variable node receives the messages \( Q_{m,n}^{[q]} \) from the neighboring check nodes and propagates back the updated messages \( Q_{m,n}^{[q+1]} \), as

\[ Q_{m,n}^{[q+1]} = \lambda_{m,n} + \sum_{n \in \Lambda_{m}^{[q]}} R_{m,n}^{[q]} \]  

(2)

- **Check node process**: the check node combines together messages \( Q_{m,n}^{[q]} \) from the neighboring variable nodes to compute the updated messages \( R_{m,n}^{[q+1]} \), which are sent back to the corresponding variable nodes. Updates are performed separately on signs and magnitudes as follows:

\[ -\text{sgn}(R_{m,n}^{[q+1]}) = \prod_{j \in (N_m \setminus n)} -\text{sgn}(Q_{m,j}^{[q]}) \]  

(4)

\[ |R_{m,n}^{[q+1]}| = \Phi^{-1} \left\{ \sum_{j \in (N_m \setminus n)} \Phi \left( |Q_{m,j}^{[q]}| \right) \right\} \]  

(5)

where

\[ \Phi(x) = \Phi^{-1}(x) = -\ln \left( \tanh \left( \frac{x}{2} \right) \right). \]  

(6)

The layered decoding scheduling improves the convergence speed and reduces the number of iteration by viewing the parity check as a sequence of check through horizontal or vertical layers. The intermediate updated messages are used in the updating of the next layer. The layered decoding principle for horizontal layers is expressed by [14].
A layered decoding architecture for LDPC Decoder with low energy consumption

Thus for a check node to compute the magnitudes of the outgoing messages, only two of the incoming messages with the smallest magnitudes have to be determined. The computation complexity of (8) is significantly reduced. Furthermore, the storage of the outgoing messages has been reduced to only two as opposed to $d_c$, where $d_c$ stands for the check node degree (i.e., the number of the neighboring variable nodes of a check node), because $d_c$-1 variable nodes share the same outgoing message value, is refreshed after every check node update. The key advantage of the layered schedule is that the immediate up-dates of the posterior messages are used and propagated to the next layers for its updating within the iteration [14], which in-creases the convergence speed and reduces the average number of iteration by up to 50% [9].

The computation of (6) and (8) are complicated and difficult for hardware implementation. Low complexity algorithms such as min-sum approximation have been proposed to reduce the computation complexity [24]. In the min-sum decoding algorithm, the computation of (8) is approximated and expressed by

$$H_{min}^{n \times k} = \min_{j \in \{1, 2, \ldots, n\}} \{T_{min}^{n \times k}(j)\}$$

Thus for a check node to compute the magnitudes of the outgoing messages, only two of the incoming messages with the smallest magnitudes have to be determined. The computation complexity of (8) is significantly reduced. Furthermore, the storage of the outgoing messages has been reduced to only two as opposed to $d_c$, where $d_c$ stands for the check node degree (i.e., the number of the neighboring variable nodes of a check node), because $d_c$-1 variable nodes share the same outgoing message value [24]. In order to achieve better performance and maintain the similar computation complexity and storage requirement of the min-sum approximation, the variants of the min-sum, such as offset min-sum [16], [17] and two-output approximation [15], have also been proposed and adopted in the hardware design.

B. The Layered Decoding Architecture

The layered decoding algorithm has been adopted in many designs [9]–[18] due to their high convergence speed and easy adaptation to the flexible LDPC codes. In this section, the defined in IEEE 802.11n decoder architecture with layered decoding algorithm for the type of the architecture-aware LDPC codes (AA-LDPC) [9] is briefly introduced. Architecture-aware codes were proposed to facilitate the hardware design of the decoder. They are structured codes, whose parity-check matrix is built according to specific patterns. They are suitable for VLSI design, because the interconnection of the decoder is regular and simple and the trade-off between throughput and hardware complexity can be easily made. Since they support an efficient partial-parallel hardware VLSI implementation, AA-LDPC codes have been adopted in several modern communication standards, such as DVB-S2 [3], IEEE 802.16e and IEEE 802.11n [5].

Fig. 2 shows an example of such parity-check matrix which is a LDPC code defined in IEEE 802.11n. It is of rate 5/6 with sub-block size (i.e., the size of the identity sub-matrix) of 81. The parity-check matrix is composed of null sub-matrix or identity sub-matrix with different cyclic shifts. The numbers stand for the cyclic shift value of the identity sub-matrix, and the “-” stands for null sub-matrix. Several VLSI architectures have been proposed for the decoder of these systems and layered decoding algorithm is commonly adopted in the design. A block diagram of these decoders is shown in Fig. 3 [13]–[19]. In the decoder, multiple soft-in-soft-out (SISO) units work in parallel to calculate multiple check node process for a layer. The Channel RAM is used to store the input LLR value of the received data initially. During the iteration of the decoding, it is used to store the posterior reliability values of the variable nodes. The shifter is used to perform the cyclic shift of the soft output messages so that the correct message is read out from the Channel RAM and sent to the corresponding SISO for calculation based on the base matrix. The Sub-array is used to perform the subtraction of (9), and the results will be sent to the SISO unit and the memory used to store these intermediate results (i.e., FIFO in Fig. 3), at the same time. The SISO unit performs the check node process of (7) and (8). The two-output approximation [15] is used for the SISO computation, and two outgoing magnitudes are generated for a check node. One is for the least reliable incoming variable node, and the other is for the rest of the variable nodes. Thus, the SISO unit, for every check node, will generate the signs for the outgoing messages of all the variable nodes, two magnitudes and an index [23]. The index is used to select the two magnitudes for the update process in the Add-array. The data generated by the SISO will be stored in the Message RAM. The Add-array

![Fig. 2. Base matrix for a rate 5/6 with sub-block size of 81 LDPC code](image-url)
performs the addition of (10), by taking the output of the SISO and the intermediate results stored in the FIFO. The results of the Add-array will be written back to the Channel RAM. To increase the throughput, pipeline operation of the decoder is adopted in the design [15], [16]. To reduce the power consumption of the Channel RAM for a layered decoding architecture, memory-bypassing scheme has been proposed in [10] by reducing the amount of memory access of the Channel RAM, which is briefly introduced in Section II-C.

C. The Memory Bypassing Scheme

In the layer decoding architecture, during the decoding, for every layer, the soft messages are read from and wrote into the Channel RAM and the FIFO every cycle. The Channel RAM stores the soft posterior reliability values of the variable nodes. The updated values are stored back to the Add-array and will be used in the decoding of the subsequent layer. In the memory-bypassing scheme [10], when two consecutive layers have non-null entry at the same column, the results of the Add-array can be directly sent to the cyclic shifter and used for the decoding of the next layer without the need of storing the intermediate result in the Channel RAM. The memory bypassing scheme saves the write operation for the current layer and the read operation for the next layer. Fig. 4 shows an example. Fig. 4(a) shows a base matrix with three layers and Fig. 4(b) shows the timing diagram of the pipeline. Without any memory bypassing, the number of read and write access of the Channel RAM is equal to the non-null entries in the matrix. In this example, the total number of read and write operation is 12. If bypassing scheme is employed, instead of writing back the channel RAM, the updated soft output values are used directly for the decoding of the next layer, the number of memory access is reduced. For example memory access for columns 0 and 1 can be bypassed for the second layer decoding and memory access for columns 0 and 3 can be bypassed for the third layer decoding. Thus, 6 out of the 12 read and write operations are eliminated, and 50% of the power consumption of the Channel RAM can be saved. Assume the pipelined architecture takes two clock cycles for the cyclic shifter, Sub-array, the SISO and the Add-array to finish the computation after the last incoming variable node is read in, a detail timing diagram showing the operation of the decoder is shown in Fig. 4(c). The order of read and write of the Channel RAM is following the natural order stated in the base matrix. Fig. 4(d) shows the memory bypassing scheme introduced in [10]. We denote it as simple memory bypassing. Here column 0 and 2 are written earlier for layer 0 and columns 0 and 2 are scheduled later for layer 1 so that overlap can be achieved. There exist some column orders that can result in maximum overlap between two layers. However, for the subsequent layers the memory overlap may not be optimized. Due to data dependency, the memory write of a certain column for the existing layer should finish before or at the same time with the reading of the same column for the subsequent layer. In order to achieve that, the decoding of the third layer has to be delayed to align the memory access and idling cycles are inserted in the decoding pipeline. We need to add idling cycle in order to maximize the overlap and even that there is still one potential overlap (W3, R3) in the third layer that cannot be achieved. From that we can see that the simple memory bypassing scheme cannot realize all the potential memory bypass operation available for the decoding matrix or extra idling cycles are needed to be inserted. In Section III, we will present an improved memory by-passing scheme which can achieve memory bypass for all the overlapped columns.

III. AN IMPROVED MEMORY BYPASSING SCHEME

In order to achieve the optimal amount of memory bypassing, we decouple the read and write order of the Channel RAM and the memory storing the intermediate messages (i.e., FIFO in the traditional design) for a layer. This is shown in Fig. 5(a). We can see that all the potential bypassings are achieved and at the same time the idling cycles are minimized. As the messages read out from the Channel RAM will be stored in the intermediate data RAM (i.e., FIFO in the traditional design) after the subtraction.
check node update, the read order of the Channel RAM is the write order of the intermediate data RAM and the read order of the intermediate data RAM is the write order of the Channel RAM. The access sequence of the intermediate data RAM is shown in Fig. 5(b). In the above example, we limited our discussion to the following assumptions.

1) Only overlapping between two consecutive layers are considered for the memory bypassing.

2) The number of the latency cycles is 2. i.e., it takes two clock cycles for the cyclic shifter, Sub-array, the SISO and the Add-array to finish the computation after the last incoming variable node is read in. To optimize the number of memory bypass, we can consider the overlapping of more layers. For example, in Fig. 4(b), the first layer and the third layer have non-null entry at column 3, and this overlapping can be used for memory bypassing. The memory operations considering the overlapping of the three consecutive layers are shown in Fig. 6. For this example, if we also consider the overlapping of the first and the third layer, two more memory access can be by-passed (i.e., the write operation W3 in the first layer and W2 in the second layer can be bypassed with the read operation R3 in the third layer and R2 in the first layer of the next decoding iteration.) Considering the overlapping of three consecutive layers, the memory-bypassing operation can be divided into two cases: memory-bypassing between layers and q+2.

Fig. 4. The memory bypassing operation for the Channel RAM in the layered LDPC decoder.

Fig. 5. Memory operations with different read and write order for the matrix shown in Fig. 4.
A layered decoding architecture for LDPC Decoder with low energy consumption

![Clock cycle diagram](image)

**Fig. 6.** Memory operations by considering the overlapping of three consecutive layers for the matrix shown in Fig. 4.

**Fig. 7.** Channel RAM operation when considering the overlapping among three consecutive layers.

Fig. 6. Memory operations by considering the overlapping of three consecutive layers for the matrix shown in Fig. 4. Fig. 7. Channel RAM operation when considering the overlapping among three consecutive layers.

$q+1$: memory-bypassing between layers $q+2$ and $q$. The potential number of memory-bypassing that can be achieved for layer $q+2$ is the number of non-null entry that the layer $q+2$ are in common with the above two layers. However, this amount of memory-bypassing may not always be achieved. The amount of memory-bypassing between layer $q+2$ and layer $q$ (i.e., case 2 shown in Fig. 7) is limited by the latency cycles. When the latency cycles are smaller than the number of non-null entry that the current layer $q+2$ are in common with the layers $q$, but not in common with layer $q+1$, some of the memory operations for the overlapped columns cannot be bypassed. In this case, to increase the amount of memory-bypassing operation, the latency clock cycles have to be increased. We limit the memory-bypassing to three consecutive layers. Since the clock node degree is in general larger than the latency cycles, the write operation of layer $q$ will finish earlier than the read operation of the layer $q+3$ unless we insert many latency clock cycles. Thus, the memory-bypassing of four or more layers is not considered.

The amount of the overlapped columns indicates the number of memory-bypassing operation that we can achieve. The amount of the overlap directly depends on the order of the layer decoding. To achieve the optimal amount of memory-bypassing, the optimal decoding order has to be determined, and also the non-null entries in every layer has to be properly scheduled such that the memory bypassing can be applied for every overlapped column. In Section IV, the algorithm for finding the optimal decoding order and the scheduling algorithm for the non-null entries inside each layer will be presented.

**IV. THE QUICK SEARCHING ALGORITHM**

To obtain the optimum decoding order, we can use brute-force method to list out all the permutation orders of the layers and then compute the number of overlapped columns of each permutation. A few experiments have been done for the example codes defined in IEEE 802.11n. The results are shown in Tables I and II which also shows the time required for the calculation. Table I shows the results when we only consider the overlapping of two consecutive layers. We also compare the number of overlapped columns obtained using different decoding orders, namely the best, the natural and the worst order. Table II shows the corresponding results when we consider the overlapping of three consecutive layers. From the results we can see that the total number of overlapped columns when considering the overlapping of three consecutive layers using the optimal order is increased by 12.9%–19.3% and the optimal decoding order and the scheduling algorithm for the non-null entries inside each layer will be presented.

| TOTAL NUMBER OF THE OVERLAPPED COLUMNS WHEN CONSIDERING THE OVERLAPPING OF TWO CONSECUTIVE LAYERS FOR THE IEEE 802.11N LDPC CODES |
|---|---|---|---|---|
| Rate | 1/2 | 2/3 | 3/4 | 5/6 |
| Order | 1/2 | 2/3 | 3/4 | 5/6 |
| Best case order | 49 | 55 | 59 | 65 |
| Natural order | 48 | 50 | 59 | 65 |
| Worst case order | 30 | 42 | 49 | 63 |

**TABLE II**

| TOTAL NUMBER OF THE OVERLAPPED COLUMNS WHEN CONSIDERING THE OVERLAPPING OF THREE CONSECUTIVE LAYERS FOR THE IEEE 802.11N LDPC CODES |
|---|---|---|---|---|
| Rate | 1/2 | 2/3 | 3/4 | 5/6 |
| Order | 1/2 | 2/3 | 3/4 | 5/6 |
| Best case order | 61 | 67 | 70 | 78 |
| Natural order | 53 | 65 | 67 | 76 |
| Time required (second) | 6850.28 | 4.02 | 0.19 | 0.64 |
memory access of the Channel RAM is reduced by the same amount.

in Fig. 2. Table II shows that for the codes with small number of the layers, the brute-force method is still practical. However, when the base matrix becomes larger, the time required for the calculation will increase dramatically. It is infeasible to use the brute-force method to find the best order of the layers for the base matrix with large number of layers. For example, the LDPC codes defined in DVB-S2 have 180 layers! The complexity of the brute-force method is and will quickly become impractical when the number increases. In order to reduce the computation time for the searching algorithm, the problem of finding the optimal decoding order of the layers which has the maximum amount of overlapping is analyzed.

A quick searching algorithm is proposed and the results for a few LDPC codes are presented. For better illustration, the algorithm finding the best order of the layers having the maximum amount of overlapped columns between two consecutive layers is considered first. We denote it as two-layer overlapping. It can be formulated as a graph problem. Let be a complete graph where and are the set of the nodes and edges of the graph, respectively. A node represents row of the base matrix and the cost of edge is the number of overlap between rows and . The problem of finding the optimal orders is the same as that of finding the path starting from any of the node in , visiting all the other nodes exactly once and returning back to the starting node and that summation of costs of the edges of the path is maximum. Fig. 8 shows the graph representing the rate 5/6 code of IEEE 802.11n. The problem is equivalent to the well-known traveling salesman problem (TSP) which is known to be NP-hard. For problems that have small number of nodes, such as the one shown in Fig. 8, brute-force approach can be used to find the optimal solution. For problem with large size, heuristic algorithm is needed to generate a good solution. The problem of finding the optimal order of the layers that considers overlapping among more than two consecutive layers is similar to the two layer overlapping problem except that the calculation of the cost function is more complicated. Hence, the computation complexity is also NP-hard and heuristic algorithm is required to find a good solution for problem with large value of . In this work, we used simulated annealing [28] to implement the heuristic algorithm for the two-layer-overlapping and three-layer-overlapping problems. The cost function used in the

<table>
<thead>
<tr>
<th>TABLE III</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOTAL NUMBER OF THE OVERLAPPED COLUMNS FOR THE IEEE 802.16E LDPC CODES</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>802.16e Sub-block size 96</th>
<th>Two layer</th>
<th>Three layer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural order</td>
<td>Solution found by simulated annealing</td>
<td>Natural order</td>
</tr>
<tr>
<td>1/2</td>
<td>23</td>
<td>38</td>
</tr>
<tr>
<td>2/3 A</td>
<td>35</td>
<td>42</td>
</tr>
<tr>
<td>2/3 B</td>
<td>9</td>
<td>52</td>
</tr>
<tr>
<td>3/4 A</td>
<td>51</td>
<td>51</td>
</tr>
<tr>
<td>3/4 B</td>
<td>55</td>
<td>62</td>
</tr>
<tr>
<td>5/6</td>
<td>67</td>
<td>67</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOTAL NUMBER OF THE OVERLAPPED COLUMNS FOR THE IEEE DVB-S2 LDPC CODES</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DVB-S2 Sub-block size 360</th>
<th>Two layer</th>
<th>Three layer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural order</td>
<td>Solution found by simulated annealing</td>
<td>Natural order</td>
</tr>
<tr>
<td>1/4</td>
<td>9</td>
<td>146</td>
</tr>
<tr>
<td>1/3</td>
<td>14</td>
<td>157</td>
</tr>
<tr>
<td>2/3</td>
<td>26</td>
<td>159</td>
</tr>
<tr>
<td>1/2</td>
<td>24</td>
<td>139</td>
</tr>
<tr>
<td>3/5</td>
<td>70</td>
<td>185</td>
</tr>
<tr>
<td>2/3</td>
<td>34</td>
<td>131</td>
</tr>
<tr>
<td>3/4</td>
<td>47</td>
<td>136</td>
</tr>
<tr>
<td>4/5</td>
<td>71</td>
<td>130</td>
</tr>
<tr>
<td>5/6</td>
<td>90</td>
<td>148</td>
</tr>
<tr>
<td>8/9</td>
<td>54</td>
<td>86</td>
</tr>
<tr>
<td>9/10</td>
<td>55</td>
<td>97</td>
</tr>
</tbody>
</table>
re-sult in significant increases in overlapping. Table III shows that for the codes used in IEEE 802.16e [4], 65.8%–98.8% of the ac-cess of the posterior reliability values in the Channel RAM can be bypassed. Table IV shows that for the codes used in DVB-S2 [3], 30.9%–66.9% of the access of the posterior reliability value in the Channel RAM can be bypassed. It can be seen that considering overlapping among more layers and using optimum decoding order have a significant effect on the increase in the number of memory bypassing. In order to implement the memory bypassing scheme, after determining the optimal decoding order of the layers, the order

of the column decoding of each layer has to be properly scheduled. Since the read and write order of the Channel RAM for a layer are decoupled, they can be scheduled differently and independently. As shown in Fig. 7, for the write order of layer , the non-null entries that the current layer are in common with the next layer are written first. The non-null entries that the layer are in common with the layer but not with the layer are written at the end. The other non-null entries are then scheduled in between randomly. From Fig. 7, we can also see that if the latency of the decoding data-path is zero, i.e., the SISO can start writing to the Channel RAM for a certain layer, e.g., layer, immediately after all the inputs are read from the Channel RAM, then the memory bypassing between layers and cannot be done. It also means that if the number of overlapping between layers and is larger than the lat- ency of the data-path, some of the potential memory bypassing cannot be realized. In this case, in order to achieve the maximum memory bypassing operation, we need to add idle clock cycles in the pipeline. When the latency is larger than the number of overlapping between layers and , memory bypassing can always be achieved for all the overlapped columns without the need of inserting idle clock cycles. In Section V, the architecture of the LDPC decoder implementing the improved memory-by-passing scheme will be presented.

V. LDPC DECODER ARCHITECTURE IMPLEMENTING THE MEMORY BYPASSING SCHEME

The block diagram of the proposed LDPC decoder for IEEE 802.11n is shown in Fig. 10. Since the sub-block size is 81,
we use 81 soft-in soft-out (SISO) units working in parallel to calculate multiple check nodes processing for a layer at the same time. The operation of the shifter, the sub-array and the SISO is the same as the traditional layered decoding architectures [13]–[15]. The received messages are quantized into 5 bit and the bit-width of the soft output for the variable nodes is 6. The two-output approximation approach [15] is used to implement the SISO. For every check node, two magnitudes, an index, and signs from all variable nodes are generated to represent the check node message and stored in the message RAM [23]. The message RAM is composed of one 88 88-bit dual port SRAM and 14 24  45-bit dual port SRAMs. The one 88 81-bit dual port SRAM is used to store the signs of the check node messages, and the 14 24 bit dual port SRAMs are used to store the magnitudes and the indexes which represent the magnitudes of the check node messages. The decoder will stop decoding if the signs of the messages in one iteration during the decoding satisfy all the parity checks or the number of iteration equals to a pre-defined value (In our design, the value is equal to 15). In order to minimize the memory access of the Channel RAM, the decoding order of the layers is determined by the algorithm introduced in the previous section. After determining the decoding order of the layers, the read and write order of the Channel RAM for the non-zero entries within a layer is scheduled to achieve the memory bypassing for all the overlapped columns and minimize the idle cycles due to the data dependency of the layers. The read and write order of the intermediate data RAM is then fixed, as they are the same as the write and read order of the Channel RAM, respectively. In the traditional decoder, the FIFO and the Channel AM are implemented separately, because of the addressing of the FIFO and Channel RAM are different. By de-coupling the read and write order of the intermediate data RAM, the intermediate data RAM and the Channel RAM can be combined because the messages stored in the same address in the Channel RAM and the intermediate data RAM not needed at the same time. When the intermediate data RAM is storing the message at the location , the message stored at the location of the Channel RAM is not needed, and vice versa. Instead of using two separate memories, we use a single four port memory (two read ports and two write ports), to implement both the Channel RAM and the intermediate data RAM. By doing so, the area of the required memory is reduced. We denote the combined memory as the new Channel RAM.

### TABLE V

<table>
<thead>
<tr>
<th>Rate</th>
<th>1/2</th>
<th>2/3</th>
<th>3/4</th>
<th>5/6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cycle per iteration</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Traditional design</td>
<td>158</td>
<td>135</td>
<td>121</td>
<td>102</td>
</tr>
<tr>
<td>Memory bypassing</td>
<td>90</td>
<td>88</td>
<td>86</td>
<td>80</td>
</tr>
<tr>
<td>Simple bypassing</td>
<td>133</td>
<td>120</td>
<td>107</td>
<td>91</td>
</tr>
<tr>
<td>Idle cycle (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Traditional</td>
<td>45.7</td>
<td>34.8</td>
<td>29.8</td>
<td>22.5</td>
</tr>
<tr>
<td>Memory bypassing</td>
<td>4.4</td>
<td>0.9</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td>Simple bypassing</td>
<td>35.3</td>
<td>26.7</td>
<td>20.6</td>
<td>13.2</td>
</tr>
</tbody>
</table>

The new Channel RAM is composed of 6 24 81 bit four port SRAMs and stores the input LLR values of the received data initially. Each entry of the new Channel RAM is dedicated to store the messages of the 81 variable nodes in the base-matrix. One pair of the read and write ports (R0 and W0) are used for the read and write access of the original Channel RAM and another pair of the read and write ports (R1 and W1) are used for the read and write access of the original intermediate data RAM. If the updated results will be used directly in the decoding of the following layers, they will be sent to the shifter directly through a mux-array. The write port W0 and the read port R0 are dis-abled. Otherwise, the updated messages will be written into the new Channel RAM through the write port W0 and the messages needed in the next layer decoding are read out through the read port R0.

A bank of muxes is added to select the output of the Add-array and that of the Channel RAM and pipeline registers are added after the Add-array, to implement the memory bypassing scheme. Because the order of the messages entering the SISO (i.e., the read order of the read port R0) and the order of the messages updated in the Add-array (i.e., the read order of the read port R1) are different, the index generated in the SISO indicating the position of the least reliable incoming messages will be incorrect for the update process. A ROM containing the order of the updated process (i.e., read order of the read port R1) is added and it is used together with the index generated in the SISO to select the two magnitudes for the update process. The additional hardware is easy to implement and the overhead in area and power is very small. By de-coupling the read and write order of the Channel RAM and using the memory bypassing scheme, the number of read and write access of the Channel RAM is reduced by 70.9%–98.7% depending on the codes.
At the same time, the idle cycles due to the data dependency of messages are minimized.

VI. EXPERIMENTAL RESULTS

We implemented the proposed LDPC decoder to demonstrate the performance of the improved memory by-passing scheme. We also implemented the traditional layered decoding architecture [14], [15] and the LDPC decoder using simple memory bypassing scheme for comparison. For all the designs, the bit-width for the soft output messages is set to 6. The decoders were implemented and synthesized with Synopsys Design Com-piler using the Artisan’s TSMC 0.18μm standard cell library. The power consumption of the embedded SRAM is characterized by hspice simulation with the TSMC 0.18μm process. The power consumption of the decoder was simulated using Synopsys VCS-MX and Prime-Time. The supply voltage is 1.8V and the clock frequency is 250MHz. The first experiment compares the clock cycle required, the idle cycles and hence the throughput of the decoder. We also compared with the design in [18], which uses the reversal order of the read operation as the order of the write operation to reduce the idle cycles. Table V summaries the results. The re-quired numbers of idle cycles reported in [18] are also included in Table V. Compared with the traditional decoder using the natural order (i.e., the order specified in the standard), the pro-posed decoder with the memory bypassing scheme decouples the read and write order of the memory and hence can reduce the number of idle cycles by 21.2%–41.3%. Compared with the simple memory bypassing scheme of which the read and write

**TABLE VI**

ENERGY CONSUMPTION (PJ/BIT/ITERATION) OF THE THREE LDPC DECODERS WHEN OPERATED IN 250 MHZ

<table>
<thead>
<tr>
<th>Message RAM</th>
<th>Rate</th>
<th>1/2</th>
<th>2/3</th>
<th>3/4</th>
<th>5/6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional design</td>
<td>16.5</td>
<td>13.1</td>
<td>10.8</td>
<td>8.7</td>
<td></td>
</tr>
<tr>
<td>Simple bypassing</td>
<td>16.5</td>
<td>13.1</td>
<td>10.8</td>
<td>8.7</td>
<td></td>
</tr>
<tr>
<td>Ours</td>
<td>16.5</td>
<td>13.1</td>
<td>10.8</td>
<td>8.7</td>
<td></td>
</tr>
<tr>
<td>Logic units</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Traditional design</td>
<td>62.3</td>
<td>54.6</td>
<td>49.4</td>
<td>43.9</td>
<td></td>
</tr>
<tr>
<td>Simple bypassing</td>
<td>60.9</td>
<td>53.0</td>
<td>48.7</td>
<td>43.9</td>
<td></td>
</tr>
<tr>
<td>Ours</td>
<td>57.7</td>
<td>50.0</td>
<td>47.5</td>
<td>43.8</td>
<td></td>
</tr>
<tr>
<td>Channel RAM + FIFO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Traditional design</td>
<td>64.7</td>
<td>66.6</td>
<td>65.5</td>
<td>62.4</td>
<td></td>
</tr>
<tr>
<td>Simple bypassing</td>
<td>47.1</td>
<td>48.5</td>
<td>43.8</td>
<td>38.6</td>
<td></td>
</tr>
<tr>
<td>Ours</td>
<td>49.4</td>
<td>40.2</td>
<td>38.4</td>
<td>32.7</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>143.5</td>
<td>134.3</td>
<td>125.7</td>
<td>115.0</td>
<td></td>
</tr>
<tr>
<td>Simple bypassing</td>
<td>124.5</td>
<td>114.4</td>
<td>103.3</td>
<td>91.2</td>
<td></td>
</tr>
<tr>
<td>Ours</td>
<td>114.6</td>
<td>103.3</td>
<td>96.8</td>
<td>85.2</td>
<td></td>
</tr>
</tbody>
</table>

**TABLE VII**

COMPARISON THE DIFFERENT LDPC DECODER IMPLEMENTATIONS

<table>
<thead>
<tr>
<th>Memory bypassing</th>
<th>Traditional</th>
<th>IEEE 802.11n</th>
<th>IEEE 802.16e</th>
<th>TDMP LDPC Reader [12]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Throughput</td>
<td>503 Mb/s</td>
<td>294.5 Mb/s</td>
<td>356 Mb/s</td>
<td>640 Mb/s</td>
</tr>
<tr>
<td>(250MHz, 10 iter, synthesis result)</td>
<td>(250MHz, 10 iter, synthesis result)</td>
<td>(240MHz, 13–14 iter, synthesis result)</td>
<td>(150MHz, 20 iter.)</td>
<td></td>
</tr>
<tr>
<td>Area</td>
<td>2.67 mm²</td>
<td>2.67 mm²</td>
<td>0.74 mm²</td>
<td>6.25 mm²</td>
</tr>
<tr>
<td>(1.26 mm² logic part)</td>
<td>(1.24 mm² logic part)</td>
<td>(0.45 mm² logic part)</td>
<td>14.3 mm²</td>
<td></td>
</tr>
<tr>
<td>Power</td>
<td>463 mW</td>
<td>339.5 mW</td>
<td>234.6 mW</td>
<td>787 mW</td>
</tr>
<tr>
<td>CMOS technology</td>
<td>180 nm, 1.8V</td>
<td>180 nm, 1.8V</td>
<td>65 nm, 1.2V</td>
<td>90 nm, 1.0V</td>
</tr>
</tbody>
</table>

order of the Channel RAM are not decoupled, the number of idle cycle is reduced by 11.9%–30.6%. Compared with the design in [18], the number of idle cycles is reduced by 1.0%–13.2%. The idle clock cycle in the decoder using the proposed memory bypassing scheme is only due to the irregular check node degrees. The idle clock cycles due to the data dependency, i.e., the up-dated message is computed before it can be used in another layer [21], [22] are all eliminated. The next experiment compares the power consumption of different decoders. Because the clock cycles required per iteration for the decoders are different, we used the energy efficiency of the decoders for comparison instead. The av-er-age energy consumptions for decoding a block of data for different code rate modes are shown in Table VI. It can be seen that the decoder using the proposed memory bypassing scheme reduces the energy consumption of the channel RAM and FIFO by 37.6%–47.5%, comparing with the traditional decoder. The corresponding reduction in energy is 27.2% to
A layered decoding architecture for LDPC Decoder with low energy consumption

38.1% when comparing with the traditional architecture using simple bypassing scheme. The overall energy reduction of the decoder is reduced by 20.1%–25.9% and 13.2%–20.7%. The reduction will be more for other LDPC codes such as DVB-S2 and IEEE 802.16e as the amount of memory bypassing that can be achieved by using the improved memory bypassing scheme is increased significantly as shown in Tables III and IV. Finally we compare the overall power consumption of the proposed decoder with other LDPC decoder implementations published in the literature. The comparisons are summarized in Table VII. Although it is hard to directly compare the different architectures since the LPDC codes are different and the implementation techniques are also different, the comparison gives some idea on the energy efficiency of different designs.

VII. CONCLUSION

I have presented an improved memory-bypassing scheme to reduce the memory access and hence the energy consumption of the LDPC decoder by exploiting the characteristic of the LDPC parity check matrix. Searching algorithm to find the optimal decoding order that results in maximum number of memory bypassing was proposed. The corresponding architecture sup-porting the proposed scheme was also presented. Experimental results show that optimum reduction in the memory access can be achieved for LDPC decoder targeting IEEE 802.11n specification and the memory access is reduced by 12.9%–19.3%. In addition, the idling cycle is reduced by 1.0%–13.2%, compared with the state-of-the-art design.

REFERENCES:

X: V-RAYS
REAL TIME ANALOGOUS X-RAY VISION TECHNOLOGY

R.VIVEKANANDAN1 & P.BARATHKUMAR2
1,2Dept. of ECE, IV Year, Sri Sairam Engineering College

Abstract-The paper throws spotlight on developing a medical instrument for complete real time analysis of bones. The process of currently available X-ray imaging technique for analysis of fracture and other bone related malfunctioning is getting tedious on faulty procedures and improper place of imaging than the required places. This costs us more time, money and energy to make a X-ray again at the proper place. The XV-Rays technology overcome the above mentioned problem and it results in developing a device that every doctor could be equipped with for real time monitoring of bones. Also the idea has been enhanced with image processing technique for analysis of types of fractures depending on the received X-ray images. This could be a pioneer in the image capturing and analysis field of bones. Also the computerised instant analysis of the images of bones would increase the efficiency of proper treatment being given to the patient and the cure is fast than ever before.

INTRODUCTION:

Today, a large number of x-ray images are interpreted in hospital and computer aided system that can perform some intelligent task and analysis is needed in order to raise the accuracy and bring down the miss rate in hospital. Conventionally, doctors in hospital examine the bone x-ray images based on their experience and knowledge whether a fracture exist. These kind of manual inspection of x-rays consume a lot of time and the process itself is Monotonous and mistakes might be made during the inspection. This study discusses the development of a system which can differentiate the fractured bone from the non-fractured bone? Other objective includes classification of specific type of fractured femur. This study will focus on femur shaft fracture detection. Femur is the longest and the strongest bone in the skeleton in our human body. It is not vertical in the erect posture and is separated above from its fellow by an interval corresponding to the breadth of the pelvis, but inclining gradually downward and medial ward, near the line of gravity of the body. The Inclination mentioned is not identical in all people and is greater in the female. Like other long bones, the femur is divisible into three main parts, which are the body (Diaphyseal) and the two extremities. The fractures of the femur's Diaphyseal will only occur during serious incident. These fractures often happen in the same extremity including the fractures of the femoral neck, posterior fracture-dislocation of the hip, tears of the collateral ligaments of the knee and osteochondral fractures involving the distal femur or patella and fractures of the tibia. It is significant to examine the joint above and the joint below fracture.

The fracture of the diaphysis of femur can be classified first into three group which are Simple, Wedge and Complex. For simple fracture, it will be classified into three groups which are spiral, oblique and transverse. For wedge fracture, it will be further classified into spiral wedge, Bending wedge, Fragmented wedge. For complex fracture, it will be classified into Spiral, segmental and irregular. Tian et al. (2003) has implemented the method of detecting femur fractures in x-ray images by computing the angle between the shaft axis and the neck axis. However this kind of method can only works on those fractures which are severe and has a significant change in the angle of the neck and shaft of the femur. This type of fracture detection using angle measurement has limitation in detecting the other type of fractures like wedge and simple. Donnelley et al. (2008) have created a CAD system for the long bone fracture detection. The system contain of four stages which are Long bone edge detection using scale-space approach, parameter approximation using the normal parameterization of the Hough transform and therefore spatially extended patterns were transformed into spatially compact features within the space of possible parameter values (ρ, θ), Diaphysis segmentation using the long-bone approximation parameters (ρi, θi)to calculate the best estimate of the bone centre-line, Fracture detection using gradient analysis. Link et al. (1997) had constructed first and higher order texture like mean and standard deviation and parameters such as apparent mineralization and total area associated with the strength regions are derived for normal and abnormal images. The statistically derived significant parameters Corresponding to the primary strength regions are fed to the neural network for training and validation.

MATERIALS AND METHODS

In this section, we describe the overall system design, image pre-processing techniques, segmentation, results analysis and classification. Figure 1 shows the block diagram of our developed algorithm. The input of X-Ray femur images, in DICOM 3.0 standard format will be interpreted into the developed...
software. The tested X-ray images were taken at 53 kV and 4 mAs and were digitized at 7 bit/pixel using a CCD camera. Based on the tags information from the header metadata, the size of the processed images are in 410×500 resolutions.

System design: Due to the nature of X-Ray image restoration, some image pre-processing techniques are necessary to eliminate the noise and image artifacts. The image pre-processing steps are including binary conversion, fine particles elimination and bone shaft detection. For the sake to ease the processes above and make it easier to be understood, the input image will be transformed into binary form for the very first. It follows by the K means based programming technique to detect the shaft area of the femur bone. Once the region of this specific area been identified, it will be cropped and recognize as region of interest for further segmentation application.

In order to detect the fracture bone, edge of bone features appears as a vital rule for the classification task. A wide conventional of edge detector have been considered, such as Sobel and Canny techniques have a shortcoming in border energies calculation, as more than two interest plotlines will be mapped within the output image. Thus, we have developed our own unique algorithm using shaft to obtain the edge. The thinnest width and thickest width will be used as the parameters for the reconstruction of region-of-interest. Lastly, the resultant region will be undergoing computerized texture analysis using GLCM, where it will be compared with threshold value in order to carry out the classification.

Image pre-processing:

Binary conversion: The image is changed to binary to ease the computing process and maximize the speed of calculating due to the rapid Boolean operators. By changing to binary, the femur shaft can be separated from the soft tissue shade which can be considered noise during the bone shaft image processing. Suppose threshold value, \( H \) within the greyscale image intensities from \( E, E+1, E+2 \ldots \) to \( F \), each pixel value is compared to \( H \) and decision is made to define a new pixel value which is either zero or one of the corresponding pixel in the output binary image.

Suppose the output image \( A \) and the input image is \( B \) and \( N \) is the image pixel array. The comparison operator is shown below:

\[
A(N) = \begin{cases} 
0 & \text{if } B(N) \leq H \\
1 & \text{if } B(N) > H 
\end{cases} 
\]

The threshold value is of extremely significance and need to be carefully selected. The quality of the image depends heavily on this threshold value.

Bone border edge detection: The bone border will be detected using Laplacian edge detector by operating a Laplacian operator convolution with the image. This involves a few steps. First the image is convolved with the Gaussian operator with the intent of blurring the image since not every edge is needed, only the border edge, then the Laplacian operator will be performed on the blurred image:

\[
F(x,y) = -\left[\nabla (G_\sigma(x,y) * I(x,y))\right]^2
\]

Where:

\[
\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}
\]

Fine particles eliminations: By using the median filter algorithm, the target pixel’s value is replaced by the median value of the neighbouring pixel in a kernel. The kernel’s is of importance in determining output image. These are the algorithm parameters that should carefully chosen depends on the image fine particles. The median filter is the filter chosen after the edge detector due to its ability to suppress isolated noise while preserving the femur bone border edge.

K-means based shaft segmentation: By using K means unsupervised clustering technique, the femur x-ray image will be clustered into two groups which are the shaft and the non-shaft area with the ‘K’ value equals to two. The purpose of this algorithm can be achieved by minimizing an objective function, in this case a squared error function, shown in Eq. (3).
Where, \((j)_{i} x - C\) is the distance measurement between the data point \((j)_{i} x\) and the cluster center \(C_{j}\), it indicates the distance difference from the data point to the cluster centre. In this study, the data point mentioned represents the number of white pixels which contain the value equal to one in each horizontal direction of the binary image. A few steps will be used in designing the algorithm. First, two cluster centre will be chosen randomly amidst the data points followed by constructing the distance difference between each data point and the cluster centre. After that, each data point will be assigned to one of the two clusters centre depends on the distance difference. When all the data points are assigned to cluster centre, the new cluster centre is then recalculated by computing the mean of each cluster group. These steps will keep repeating itself until the cluster centre no longer change. By then, the objective function reaches its minimum value.

Grey level co-occurrence matrices: Since the GLCM were proposed by Haralicke (1979), it has been utilized as the main tool in image texture analysis. Haralick suggested statistics equations that can be calculated from the co-occurrence matrix and be used in describing the image texture. It is a statistical way to indicate image texture structure by statistically sampling the pattern of the grey-levels occurs in relation to other grey levels. There are mostly weighted averages of the normalized co-occurrence matrix contents by multiplying a weighted average multiplier with the intent of expressing the relative significance of the value. Figure 2 shows the fundamental of \(N \times N\) matrix. The matrix shown above is squared with dimension \(N\) which represents the number of gray levels in the image. Element \([i, j]\) of the matrix is generated by computing the frequency of a single pixel with value \(i\) is adjacent to the pixel with value \(j\) and then divide the matrix by the total number of such comparisons made. Each value in the matrix is hence become the probability that a pixel with value \(i\) which will be found adjacent to a pixel of value \(j\). The mentioned co-occurrence matrix features formula which computed from the Image \((i, j)\) are.

Angular second moment:

\[
\text{ASM} = \sum_{i,j=1}^{N-1} P_{i,j} \quad (4)
\]

The image energy is the square root of ASM which is a very useful second order texture statistical operator that able to measure the uniformity of the targeted window from image. Each element in the co-occurrence matrix in Fig. 2 is squared and summed up. If in the case of a uniform image which has been normalized and all the pixels values equals to ‘1’, the output of this operator is equals to ‘1’. If the uniform image is not normalized, the output value equals to the number of the window’s size. Recall that Normalization changes the number of combination into a probability value range from zero to one. In Fig. 2, all the value has been normalized, therefore the element is \(p(i,j)\) which is a probability:

\[
\text{Energy} = \sqrt{\sum_{i,j=1}^{N} p_{i,j}^2} \quad (5)
\]

\[
\text{contrast} = \sum_{i,j=1}^{N} p_{i,j} (1 - |i - j|) \quad (6)
\]

The diagonal elements of the matrix shown in Fig. 2 indicates the group of pixel which has no difference in pixel value. For instant, the first element \(p(1, 1)\), shows that the probability of the combination of two pixels in certain spatial relation which are of the same value equals to ‘1’, the second element represents the probability of the combination of two pixels in certain spatial relation which are of the same value equals to ‘2’ and so on. In other words, it shows no contrast in diagonal elements in the co occurrence matrix. Therefore if the intent is to measure the contrast of the texture which the matrix represented, weights need to be created so that the output value of Eq. 4 shows larger value accordingly with the degree of contrast with the fact that the contrast increases as the distance of elements in the matrix from the diagonal increases. For example the element \(p(1, N)\) and \(p(N, 1)\) should show much lower value of contrast compared to the element \(p(1, 3), p(2, 2)\) in Fig. 2.

\[
\text{Dissimilarity (DIS)} = \sum_{i,j=1}^{N} p_{i,j} (i - j)^2 \quad (7)
\]

\[
\text{GLCM correlation} = \sum_{i,j=1}^{N} P_{i,j} \left[ \frac{(i - \mu_i)(j - \mu_j)}{\sigma_i \sigma_j} \right] \quad (8)
\]

The GLCM correlation indicates the linear dependency of the linear dependency between the grey levels and the neighbouring pixels where \(\mu_i\) represents the horizontal mean in the matrix, \(\mu_j\) represents the vertical mean in the matrix, \(2 \sigma_i\) and \(2 \sigma_j\) represents dispersion around the mean of combinations of target and neighbour pixel:

\[
\text{GLCM \mu}_{\text{horizontal}} = \mu_i = \sum_{i,j=1}^{N} i p_{i,j} \quad (9)
\]

\[
\text{GLCM \mu}_{\text{vertical}} = \mu_j = \sum_{i,j=1}^{N} j p_{i,j} \quad (10)
\]

\[
\text{GLCM Variance} = \sigma_i^2 = \sum_{i,j=1}^{N} (i - \mu_i)^2 p_{i,j} \quad \text{and} \quad \sigma_j^2 = \sum_{i,j=1}^{N} (j - \mu_j)^2 p_{i,j} \quad (11)
\]

\[
\text{Entropy (ENT)} = \sum_{i,j=1}^{N} \ln p_{i,j} \quad \text{and} \quad \text{Homogeneity (HOM)} = \sum_{i,j=1}^{N} \frac{p_{i,j}}{(1 + (i - j)^2)} \quad (12)
\]

where, \(\mu_i, \mu_j, \sigma_i, \sigma_j\) are the means and Standard Deviations. Homogeneity, also called ‘Inverse
Difference Moment’ is an inversion to the contrast. In calculating the contrast, the weight of element increases when distance of element from diagonal increases. Inversely, while calculating the Homogeneity, the weight of element decreases as the distance of elements from diagonal increases. In short, the weight of contrast \((4)\) is \((i-j)^2\), on the other hand, the weight of Homogeneity is \(1/(1+(i-j)^2)\).

**EXPERIMENTATION & RESULTS:**

Figure 3 shows part of our experimental results using sample patients’ data and the obtained findings demonstrated that the state of the art of computerized GLCM method is able to produce accurate border in most of the samples. Figure 4 shows the resultant image conversion into binary form, where the value 0 represents black pixel and value 1 represents the white pixel in the image. Figure 5 shows that a blue line is drawn to separate the shaft and epiphysis of femur. The algorithm enable each horizontal line of the picture inspects and identifies the white pixel among each row of pixels. When the white pixels of certain horizontal line exceed a threshold value, it will sketch a line to distinguish the shaft and the upper epiphysis of femur. After filtering the image, the image will be processed to get the edge and a red dot will be drawn on each point of edge detected as shown in Fig. 7. After that, the middle point the shaft will be calculated using the average of the width of the thinnest part of the shaft and the thickest part of the shaft. After obtaining the middle point, each middle point will be extended to left and right according to the shaft size as shown in Fig. 8. The areas within both blue lines represent the region-of-interest of our later GLCM measurement.

**GLCM calculation of fractured and non-fractured femur:** The region-of-interest of the femur x-image contain Fractures will be the region will undergo GLCM for every 50 pixel length and the length depends on the width of the size of shaft. The area will be scanned and obtain the GLCM values, for the image that 500×400. For each area, there will be four statistical GLCM values calculated in four directions which are \([1,0]\), \([0,1]\), \([1,1]\), \([-1,-1]\). Only the average of the four directions will be taken and therefore for every area there will be total 4 values. Energy provides the sum of squared elements in the GLCM also known as uniformity or the angular second moment. The values of energy indicate the texture of the area, if it is equal to 1, it means the area is uniform and no fracture found and vice versa. Contrast measures the local variations in the gray-level co-occurrence matrix. Correlation measures the joint probability occurrence of the specified pixel pairs. Homogeneity measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal.
After obtaining the statistical value of GLCM, four statistical values which are contrast, correlation, Energy, Homogeneity will be analyzed. The algorithm itself will calculate the value compare and classified automatically and indicate to the user the area of fracture if any exist. In order to access the performance and usefulness of the developed system in a real application, a thorough evaluation of the method was carried out at the Biomedical Research Centre, University Technology Malaysia, Malaysia. We ran the algorithm on a set of X-Ray images (n = 30), with 410×500 sized resolutions for performance evaluation. The images were catalogued into two different testing groups of k1 and k2 respectively. Each group of testing catalogue consisted of 15 numbers of X-Ray images. The first group k1 were images shows normal femur bone screening, where the second group of images k2 were femur bone images with fracture found. Table 1 lists the performance of developed software on k1-k2 groups of images. Simulations result shows that the developed algorithm is capable achieving as high as accuracy about 86.67 percent and able to provide reliable and consistent findings.

Based on Eq. 13-16 the calculated sensitivity was 80 percent (12 out of 15), the specificity was 93.33 percent (14 out of 15), the positive predictive value was 92.31 percent (12 out of 13) and the negative predictive value was 82.35 percent (14 out of 17). These results indicate the developed diagnostic model making well recognition and detection of bone fracture:

| Table 1: Performance of developed algorithm for bone fracture classification |
|-----------------------------|-----------------------------|-----------------------------|
| Threshold        | Group k1 | Group k2 | Accuracy      |
| Threshold > 0.95 | 12 True-Positive (TP) 1 False-Positive (FP) | 36.67% |
| Threshold < 0.95 | 3 False-Negative (FN) | 14 True-Negative (TN) | 89.5% |
| Total           | 12 15    | 15 15    | 86.67%        |

Consecutively, the real time x-ray visual system would be a great benefit for easy and fast detection of fractures by the doctors themselves with a portable sort of device and the image processing technique would enhance the accuracy of the patient getting the proper treatment and fast cure. We have proposed a method for automated femur bone fracture detection using GLCM computerized techniques. From this method we are able to classify the absence and presence of bone fracture based on the obtained parameter value from GLCM value. The threshold bordering the absence and presence of bone fracture is set to a value of 0.95. The accuracy of the developed algorithm is achieved at least 86.67 percent which promises an efficient method to Recognize bone fracture automatically. Findings show that the system is able to provide consistent and reproducible result

REFERENCES


