**TABLE OF CONTENTS**

|  |  |  |
| --- | --- | --- |
| **CHAPTERNO.** | **TOPIC** | **PAGE NO.** |
|  | Certificate from Supervisor | **2** |
|  | Acknowledgement | **3** |
|  | Summary | **4** |
| **1.** | Introduction | **7** |
| **2.** | Software design for processing | **11** |
| **3.** | Character Extraction | **14** |
| **4.** | Pattern recognition and excel sheet generation | **24** |
| **5.** | Hardware Design | **29** |
| **6.** | Challenges faced | **33** |
| **7.** | Future scope |  |
|  | REFERENCES | **49** |
|  |  |  |
|  |  |  |

**CERTIFICATE**

This is to certify that the work titled “**EXAM PAPER READER”** submitted by “**DEEPANSHARORA (7102257), DEEPAK (7102198) and VISHNU TEJA (7102289)**” in partial fulfilment for the award of degree of bachelor of technology of Jaypee Institute of Information Technology University, Noida has been carried out under my supervision. This work has not been submitted partially or wholly to any other University or Institute for the award of this or any other degree or diploma.

Signature of Supervisor:

Name of Supervisor: **Mr. Naveen Babu**

Designation: **Senior lecturer**

Date:

**ACKNOWLEDGEMENT**

We wish to express our sincere gratitude to M**r. Naveen Babu**, senior lecturer, Department of ECE, Jaypee Institute of Information Technology University, Noida for providing us an opportunity to do this project work on “**Exam Paper Reader**” .This project bears an imprint of many research works and Journals. We sincerely thank him as a mentor who was always there for guidance and encouragement in carrying out this project work.

We also wish to express our deep gratitude to the officials and other staff members of ECE Department who rendered their help by providing a good environment throughout the period of this project work.

.

**Yours faithfully:**

DEEPANSH ARORA (7102257) Batch-A4

DEEPAK SHARMA (7102198) Batch-A2

VISHNU TEJA .Y (7102289) Batch-A6

**SUMMARY**

Our Project, Exam Paper Reader also called as EPR has been designed with the motive of reducing the burden on faculty of JIIT. The main task of EPR is extracting the Enrollment number & name written by the student and corresponding marks assigned by teachers in the answer sheet. In this Optical character recognition (OCR) technology is used to extract the text written by Teachers and students. A hardware was also designed to process the bundle of papers and make the task easier.

There are three essential elements to OCR technology—scanning, recognition, and reading text. Initially, a printed document is scanned by a camera. OCR software then converts the images into recognized characters and words. The information is stored in an electronic form, either in a personal computer (PC) or the memory of the OCR system (which have IC APR9600) itself.

All OCR systems create temporary files containing the texts' characters and page layout. In our OCR's these temporary files will be converted into formats retrievable by commonly used computer software such as word processors and spreadsheet and database software. The text that is extracted after the processing is done is stored in an Excel sheet. This helps in updating the marks.

Special cameras and pattern recognition programs have been used for recognizing specially designed letters and numbers, such as the account numbers on checks. The camera converts the pattern of each letter into a binary code. A computer is programmed to process the binary code and determine which letter it represents.

In the last few years, devices and programs have been developed which make it possible for computers to recognize most typewritten characters and to adjust automatically for different type styles and sizes. In the next few years, this technology is likely to be perfected and become more widely available. To make the extraction more efficient we have also worked on the extraction of handwritten letters. As the hand written letters varies from a person to another we have collected the templates of different students and by the method of correlation the data has been extracted. (Only very limited success can be expected with handwritten letters, due to the large variations found in even one person's handwriting.).

**CHAPTER 1**

**INTRODUCTION**

Our project is divided into 3 phases-

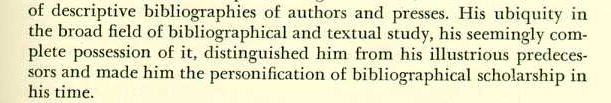
1. Optical character recognition

2. Designing a hard ware that process the bundle of papers

3. Integration of above 2 to develop a sound application

**First** we define optical character recognition (OCR) whose goal is to classify optical patterns (often contained in a digital image) corresponding to alphanumeric or other characters. The process of OCR involves several steps including segmentation, feature extraction, and classification. Each of these steps is a field unto itself, and is implemented in MATLAB. OCR is mechanical or electronic translation of images of handwritten, typewritten or printed text (captured by a scanner or webcam) into machine-editable text. We will use text extraction to convert paper books and documents into electronic files, for instance, to computerize an old record-keeping system in an office or to help blind man to dictate particular page he want to read.

One example of OCR is shown below. A portion of a scanned image of text, borrowed from the web, is shown along with the corresponding (human recognized) characters from that text



of descriptive bibliographies of authors and presses. His ubiquity in the broad field of bibliographical and textual study, his seemingly com-plete possession of it, distinguished him from his illustrious predeces-sors and made him the personification of bibliographical scholarship in his time.

Figure 1.1: Text extracted from scanned image

A few examples of OCR applications are listed here. The most common for use OCR is the first item; people often wish to convert text documents to some sort of digital representation.

1. People wish to scan in a document and have the text of that document available in a word processor.
2. Recognizing license plate numbers
3. Post Office needs to recognize zip-codes
4. Extracting data from attendance sheets and updating the earlier records.

**Second,** Designing a device that can process the papers one by one from a lot .To achieve this we have initially started our work by modifying a HP printer. The disadvantage of the printers available in the market is that they get jammed whenever the thickness of paper varies, usually the answer sheets in JIIT consists of 8 papers. Selecting a bundle of 8 papers from a lot is a little tough task. To have an efficient processing of papers we have designed a hardware which has both rack & pin mechanism and Conveyor belt system. The images are taken by the webcam which is mounted above the conveyor belt .To maintain the constant intensity of light the whole device is closed by a box with a light placed in it. In institutions faculty do change every year, so interaction to the new faculty can be done by providing a proper instructions. Most of the instructions to any device are provided in written form which requires a lot of attention to understand. In order to overcome this we have added a voice module which can record and play the messages accordingly.APR9600 module has been selected as it can store 8 messages with a time period of 60 sec.

Finally interaction between webcam and hardware is utmost important. This synchronisation was attained by serial communication between microcontroller and computer.

###### OCR converts a graphical image of a document to word converts graphical image of a document to a word processor file without typing the characters. In other words, you can change printed text into an editable, searchable document. Prior to OCR technology human was forced to type all the information available on the paper in printed format for applications such as mailing, correction etc.

###### 2.1 The Classification Process

(Classification in general for any type of classifier) There are two steps in building a classifier-training and testing.

These steps can be broken down further into sub-steps.

1. **Training**
   1. Pre-processing – Processes the data so it is in a suitable form.
   2. Feature extraction – Reduce the amount of data by extracting *relevant* information—usually results in a vector of scalar values.

(We also need to NORMALIZE the features for distance measurements)

* 1. Model Estimation – from the finite set of feature vectors, need to estimate a model (usually statistical) for each class of the training data

1. **Testing**
   1. Pre-processing
   2. Feature extraction – (both same as above)
   3. Classification – Compare feature vectors to the various models and find the closest match. One can use a distance measure.

### Pre-processing

#### 2. Recognition

**(Testing)**

#### 1. Training

**Training**

**Data**

### Pre-processing

**Test**

**Data**

**Feature**

**Extraction**

**Feature**

**Extraction**

**Classification**

**Model**

**Estimation**

## Figure 2: The pattern classification process

**CHAPTER 2**

**SOFTWARE DESIGN**

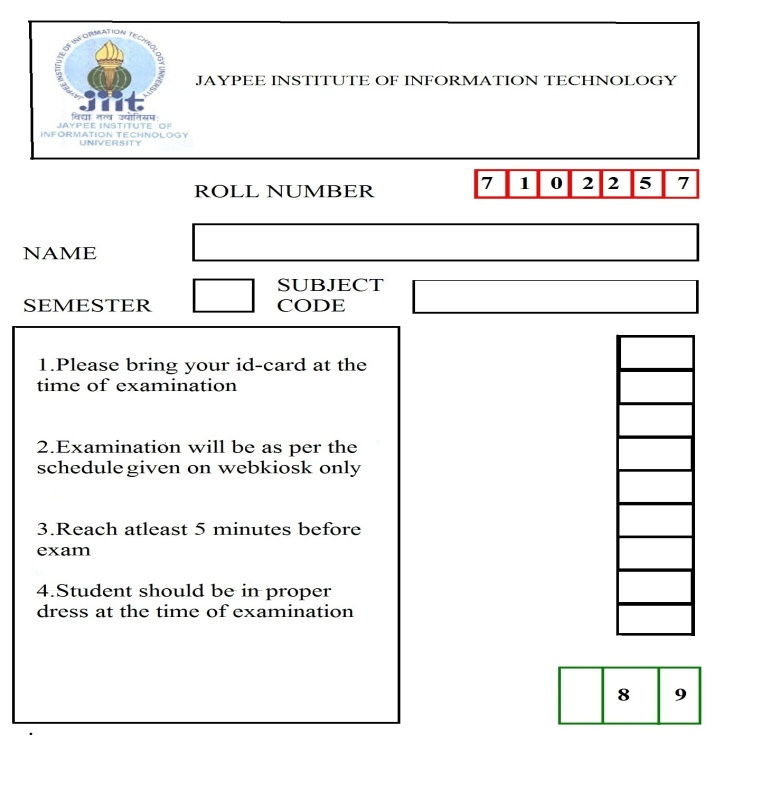
**Region Extraction:-**

In purposed prototype of exam answer script various information will be filed by students. These information will be filed in specific region of the page. Now after taking images through purposed automatic EPR machine, different region of information have to extract from the camera images. Later the extracted region will converted to a new image, the size of this new image is much less than the previous image so this will increase the processing speed for further algorithms.

Different steps of our project described below deals with the finding of numerals of roll number or marks from the mapped original image containing box. The following flowchart of sequence steps deals with the preprocessing of the image. Though the image has been taken under isolated environment provided by the hardware and in proper lighting condition, but the need of preprocessing is certain for region extraction. The camera used is also VGA quality so preprocessing becomes important step in text extraction in real images. The application of the following steps applied definitely improves the quality of image as needed for recognition by the OCR engine.

The prototype has been suggested by us as shown below which contains the region having **enrollment number** and **marks** in different color boxes to which color processing can be easily applied for the extraction of the region of interest (roi).so, we are dealing with two regions as can be seen in below prototype model:

**Roll number block**



**Marks block**

Fig2.1: Suggested prototype

**Pre-processing using CIE criteria (chromaticity):-**

As shown in fig.(2.1) the region is bounded by red color and green color boxes. Red and green both are primary color and these colors have the maximum diversity from all other colors. Blue color is too a primary color. This could be understood through the CIE (chromaticity) graph of colors.

The diagram presents visible colors as a function of *x* (red) and *y* (green) components called *chromaticity coordinates*. Positions of various spectrum colors (from violet to red) are indicated as the points of a tongue-shaped curve called *spectrum locus*. The straight line connecting the ends of the curve is called the *purple line*. The point of equal energy represents the CIE standard for white light. Any point within the diagram represents some mixture of spectrum colors. The pure or fully saturated colors lie on the spectrum locus. A straight-line segment joining any two points in the diagram defines all color variations that can be obtained by additively combining these two colors. A triangle with vertices at any three points determines the gamut of colors that can be obtained by combining corresponding three colors. The structure of the human eye that distinguishes three different stimuli, establishes the three-dimensional nature of color.

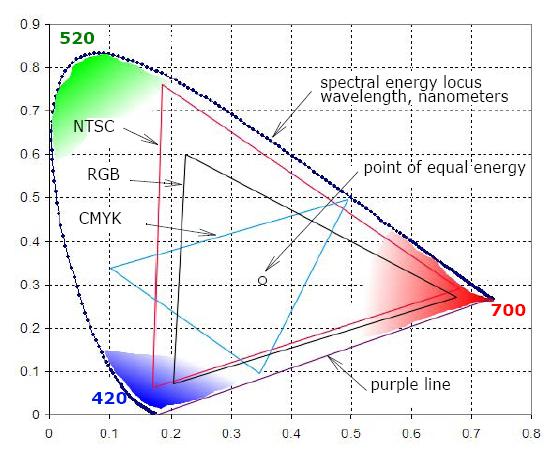
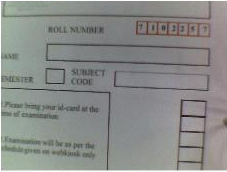


Fig 2.2 chromaticity diagram

In practice the images are taken through the web cam mounted in EPR machine and at the real time the quality of images are poor in such images.



Fig(2.3): original image taken for camera

**Software design of region extraction:**

Image taken by camera will undergo following processing steps:

CONNECTED COMPONENTS EVALUATION

MORPHOLOGICAL

PROCESSING

BINARIZARTION

HIT N MISS TRANSFORM

MEDIAN FILTERING

MORPHOLOGICAL

PROCESSING

MEDIAN FILTERING

SKELETONIZATION

NOISE REMOVAL USING AREA OF CONNECTED COMPONENTS

PROCESSED IMAGE PASSED TO OCR ENGINE

EXTREME COORDINATES

MAPPING OF EXTREME COORDINATES TO ORIGINAL RGB IMAGE

GRAY LEVEL CONVERSION

HSI CONVERSION

SATURATION LEVEL RECOGINATION

SUBTRACTION OF SATURATION LEVEL AND GRAY LEVEL

RGB SMOOTHING

DIVERGENCE OF IMAGE W.R.T RED COLOR

SHARPENING

**2.1 RGB Smoothing:**

Relative divergence calculation of RGB image:-

In RGB color mode all the color is derived trough the relative amount of three primary colors Red Green Blue. Now at macro level for any particular color there is number possible ratio set in these primary colors. Especially due to white noise the range of colors enrages. The relation between eyes, camera noise, and environment parameter couldn’t derive. For a robust system we try to obtain the relative divergence of color ratio and after a number of iteration we obtain the boundary condition for particular system.

Relative Divergence of color =

((red component – green component) + (red component – blue component))/2

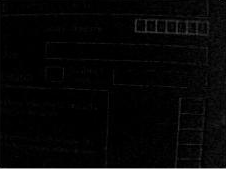


Fig 2.3 relative divergence w.r.t to red color

(b)Sharpness enhancement:-

The principal objective of sharpening is to highlight fine detail in an image or to enhance detail that has been blurred, either in error or as a natural effect of a particular method of image acquisition.

Result obtains by previous treatment passed through the sharpness filter for emphasizing demonotonousness of image.

Derivation of sharpness filter:-

Two-dimensional second order derivatives for image enhancement. The approach basically consists of defining a discrete formulation of the second-order derivative and then constructing a filter mask based on that formulation. We are interested in *isotropic* filters, whose response is independent of the direction of the discontinuities in the image to which the filter is applied.

Fig2.4 mask of sharpness filter

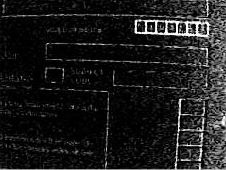


Fig 2.5 After sharpening

**2.3.Binarization of Image:-**

As the region of interest is much brighter than the rest of image so an optimized threshold could be set for the process and a binary image for further processing should be developed. also since the gray level image is more complex than binary image so binarization is needed to remove complexity during programming



Fig 2.6 binarization of Image

**2.4 Hit and Miss transform:-**

Hit and miss transform fell in the category of morphological image processing, it is a technique of identify particular shape in an image. All the morphological image processing is done on the binary image.

Definition of hit and miss transform:-

Here: - A is the input image

B is the output image

C is the stretching sub image for hit

D is the stretching sub image for miss

Using this process we have recognize the unwanted dotes and broken pieces of lines and with this information we remove these points from original image and obtain much clear image.

**2.5 Median filtering:**

The Median Filter block replaces the central value of an M-by-N neighborhood with its median value. If the neighborhood has a center element, the block places the median value there, as illustrated in the following figure-



Fig2.7 Median filtering pixel replacement pattern

The median filter is normally used to http://homepages.inf.ed.ac.uk/rbf/HIPR2/mote.gifreduce noise in an image, somewhat like the mean filter. However, it often does a better job than the mean filter of preserving useful detail in the image.

By calculating the median value of a neighbourhood rather than the mean filter, the median filter has two main advantages over the mean filter:

* The median is a more robust average than the mean and so a single very unrepresentative pixel in a neighbourhood will not affect the median value significantly.
* Since the median value must actually be the value of one of the pixels in the neighbourhood, the median filter does not create new unrealistic pixel values when the filter straddles an edge. For this reason the median filter is much better at preserving sharp edges than the mean filter.

The main work of median filter here is to deal with salt and pepper type noise.

Connected components:

Connected components labeling scans an image and groups its pixels into components based on pixel connectivity, i.e. all pixels in a connected component share similar pixel intensity values and are in some way connected with each other. Once all groups have been determined, each pixel is labeled with a gray level or a color (color labeling) according to the component it was assigned to.

Algorithm for connected componets-

For each pixel on the line, this implementation first checks to see if the pixel to the left has the same pixel value. If so, we know for sure that we're in the same blob, and the current pixel is so labelled. If the pixel at the top has the same value as the pixel to the left but not the same blob ID, we know at once that the pixels to the left and to the top are in the same region and that these regions should be merged.

If a pixel is found with left and top neighbour having different pixel values, a new blob id is created and assigned to this pixel.

The algorithm continues this way, creating new blobs and merging them whenever they are discovered to be the same. The key to a fast algorithm, however, is how this merging is done. This algorithm uses the union-find data structure which provides excellent performance for keeping track of equivalence relationships. Union-find essentially stores labels which correspond to the same blob in a tree, making it easy to remember the equivalence of two labels.

Then counting number of pixels for each label corresponds to the area, and the area having larger number of pixels corresponds to the desired area which needs to be extracted from original image

**2.6 Implementation of morphological processing:-**

It includes erosion, dilation and skeletization of image defined as below-

Erosion:-

erosion is tool of morphological image processing, as the name indicate it remove the external corner of comports in the image. Thus after this all the strongly connected component remains rest disappears.

Definition:-

A \ominus B = \{z\in E | B_{z} \subseteq A\}

B_z = \{b+z|b\in B\}\forall z\in E

Where *Bz* is the translation of *B* by the vector z, i.e., , .

And A is the image

Example:-

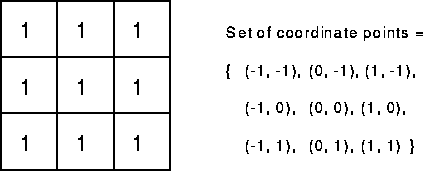


Fig2.8 sample element for erosion opration

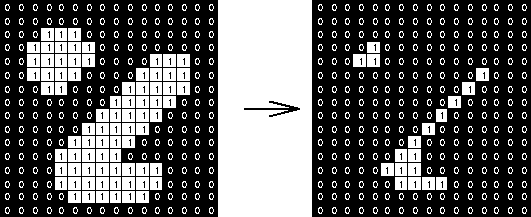


Fig sample image Fig after erosion with sample element

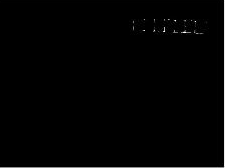


Fig 2.10 After erosion

Dilation:-Dilation is the reverse process of the erosion because all the unwanted parts are removed from the image now we can expand the region of extraction.

Example:-

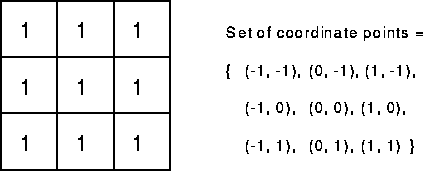


Fig Sample element for dilation

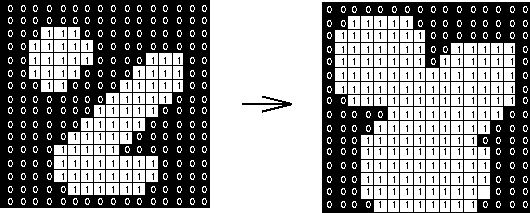


Fig sample image Fig after dilation



Fig 2.11 Output after dilation

Now as the final step we will calculate the extreme diagonal coordinates of the recognized region and using these we will extract a new image from the original RGB image. And store it as a new image.

In the next processing step this image will be send to perform next operation in order to extract the numerical value.



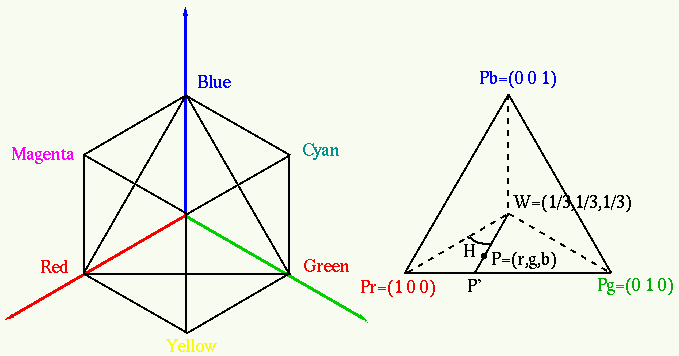
Fig 2.12 Sub image after coordinate extraction

**2.7 RGB to HSI model conversion:-**

In order to remove the red box from the enrolment number with higher aqurancy we require to convert the image in its HSI equitant model for perfectly removal of red color as in our previous attempt we had calculated relative divergence but as results was not up to a marks from where a robust system can be synthesizes for EPR machine. In HIS model we calculate the three parameter hue, intensity and saturation as saturation inform us about the primary nature of the color. Black and white both are a non primary color so they will be easily distinguished by the red with the information of saturation level of colors.

Derivation of HSI model:-

Given the intensities of the three primaries RGB of a color, we can find its HSV representation using different models. Here we use the RGB plane of the cube to find the corresponding HSV. The three vertices are represented by Pr, Pg and Pb the three components of the given color is represented by a 3D point P=(R G B). We also assume the intensities are normalized so that the R,G and B values are between 0 and 1, so that point P is inside or on the surface of the color cube.



* Determine the intensity I:

One of the definitions of intensity is I=1/3(R+G+B)

* Determine the hue H:

First find the intersection of the color vector (R G B) with the RGB triangle:

R+G+B=1

\begin{displaymath}\left\{ \begin{array}{l}
r\stackrel{\triangle}{=}R/(R+G+B)=R...
...
b\stackrel{\triangle}{=}B/(R+G+B)=B/3I
\end{array} \right. \end{displaymath}

This point P=(r, g, b) is on the RGB triangle as r + g +b =1. Here we assume the point p is inside the triangle formed by points w, Pr, and Pg. The hue is the angle <H formed by the vectors pw and Prw . Consider the dot product of these two vectors:

\begin{displaymath}\overline{P_r w} \cdot \overline{pw}
=(P_r-w)\cdot(p-w)=\vert P_r-w\vert\;\vert p-w\vert\;cos \angle H \end{displaymath}

$P_r=(1,0,0)$ $p=(r,g,b)$ $w=(1/3,1/3,1/3)$

\begin{displaymath}(P_r-w)=(1,0,0)-(1/3,1/3,1/3)=(2/3, -1/3, -1/3) \end{displaymath}

\begin{displaymath}(p-w)=(r-1/3, g-1/3, b-1/3) \end{displaymath}

\begin{displaymath}
(P_r-w)\cdot(p-w)=\frac{2}{3}(r-\frac{1}{3})
-\frac{1}{3}(g...
...)-\frac{1}{3}(b-\frac{1}{3})
=2r-g-b=\frac{2R-G-B}{3(R+G+B)}
\end{displaymath}

\begin{displaymath}\vert P_r-w\vert=\sqrt{(1-\frac{1}{3})^2+(0-\frac{1}{3})^2+(0-\frac{1}{3})^2}
=\sqrt{\frac{2}{3}} \end{displaymath}

|  |  |  |  |
| --- | --- | --- | --- |
| $\displaystyle \vert p-w\vert$ | $\textstyle =$ | $\displaystyle \sqrt{(r-\frac{1}{3})^2+(g-\frac{1}{3})^2+(b-\frac{1}{3})^2} =\sqrt{\frac{9(R^2+G^2+B^2)-3(R+G+B)^2}{9(R+G+B)^2}}$ |  |
|  | $\textstyle =$ | $\displaystyle \frac{\sqrt{6(R^2+G^2+B^2-RG-GB-BR)}}{3(R+G+B)}$ |  |

Now the hue angle can be found to be

|  |  |  |  |
| --- | --- | --- | --- |
| $\displaystyle \angle H$ | $\textstyle =$ | $\displaystyle cos^{-1}[\frac{(P_r-w)\cdot(p-w)}{\vert P_r-w\vert\;\vert p-w\vert}] =cos^{-1}[\frac{2R-G-B}{\sqrt{2/3}\sqrt{6(R^2+G^2+B^2-RG-GB-BR)}}]$ |  |
|  | $\textstyle =$ | $\displaystyle cos^{-1}[\frac{3R-(R+G+B)}{2\sqrt{R^2+G^2+B^2-RG-GB-BR}}] =cos^{-1}[\frac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}}]$ |  |

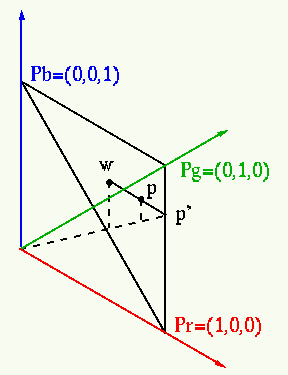
If

$B>G$

Then

$\angle H=360-\angle H$.

Determining S:



The saturation of the colors on any of the three edges of the RGB triangle is defined as 1 (100% saturated), and the saturation of w=(1/3,1/3,1/3) is zero. Denote as p’ the p=w intersection of the extension of line wp with the edge. If the normalized color is, Sp=0, and if p=p’, Sp=1. The saturation of any color point p between w and p’ is defined as

\begin{displaymath}S_p=\frac{\vert wp\vert}{\vert wp'\vert}
=\frac{\vert wp'\ve...
...1-\frac{\vert pp'\vert}{\vert wp'\vert}
=1-\frac{b}{1/3}=1-3b \end{displaymath}

Here it is assumed that point p is inside the triangle PrwPg so that b=min(r,g,b). In general

\begin{displaymath}S_p=1-3\;min(r,g,b)
=\left\{ \begin{array}{cl}
0, & min(r,g,...
...the edges} \\
0<S_p<1, & 0<min(r,g,b)<1/3 \end{array} \right. \end{displaymath}



Fig 2.13 HIS Model

Fig 2.14 hue component Fig 2.15 intensity component Fig 2.16 saturation component Fig 2.14 hue component Fig 2.15 intensity component Fig 2.16 saturation component

**2.8 RGB to grey scale conversion:-**

In Image Processing conversion of images is often are requirement. For instance, we can reduce complexity of our processing by converting our RGB image to Gray-Scale image.

Y = 0.3\*R + 0.59\*G + 0.11\*B, i.e. the out image is the result of 30% of red component, 59% of green and 11% of blue.

A two dimensional matrix having values between 0 and 255. ‘0’ representing pure black and ‘255’ representing pure white and in between them, the shades of gray.

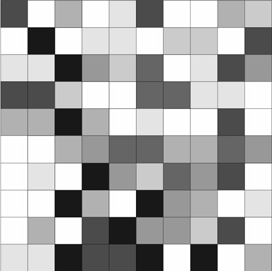


Fig 2.17 gray level image



Fig2.18 gray level image

**2.9 Calculation of histogram equivalent:**

To transfer the gray levels so that the histogram of the resulting image is equalized to be a constant:

The purposes of the solution:

* to equally use all available gray levels;
* for further histogram specification.

Histogram modeling techniques (*e.g.* histogram equalization) provide a sophisticated method for modifying the dynamic range and contrast of an image by altering that image such that its intensity histogram has a desired shape. Unlike contrast stretching, histogram modeling operators may employ *non-linear* and *non-monotonic* transfer functions to map between pixel intensity values in the input and output images. Histogram equalization employs a monotonic, non-linear mapping which re-assigns the intensity values of pixels in the input image such that the output image contains a uniform distribution of intensities (*i.e.* a flat histogram). This technique is used in image comparison processes (because it is effective in detail enhancement) and in the correction of non-linear effects introduced by, say, a digitizer or display system.

**2.10 Skeletonization of image:**

|  |  |
| --- | --- |
| http://cgm.cs.mcgill.ca/~godfried/teaching/projects97/azar/thick_t.gif | http://cgm.cs.mcgill.ca/~godfried/teaching/projects97/azar/thin_t.gif |

*Skeletonization* is the process of peeling off of a pattern as many pixels as possible without affecting the general shape of the pattern. In other words, after pixels have been peeled off, the pattern should still be recognized. The skeleton hence obtained must have the following properties:

 As thin as possible

 connected

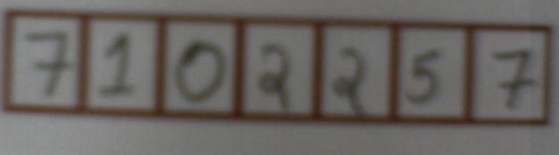
 centered

Now the image after such steps obtained and is shown in their perspective heading. The image obtained at the end of the following processing is the numerals image which can be now passed to the optical character recognition function which recognises the letter and pass the letter obtained to be saved in excel sheet along with its corresponding numbers allotted in a particular subject.



FIG 2.19 Skeletonisation Fig 2.20 Inversion of Skeletonised Image

Different Output Images while undergoing processing in function:-















**CHAPTER 3**

**CHARACTER EXTRACTION**

In this part, the characters will be identified as letters and the image will sent for recognition and further processing. After the image is cleaned up becomes a binary image which contains only the text, the binary image is then saved and the memory is cleaned up. This step is very important to increase the speed of system. After that the following steps should be done.

**5.1** **Removal the borders**

The borders have been removed; this will reduce the image size. Only the rectangular part of the image which contains the text will remain. As shown in figure, the gray region will be removed which will make the image size smaller and consequently make the program faster.

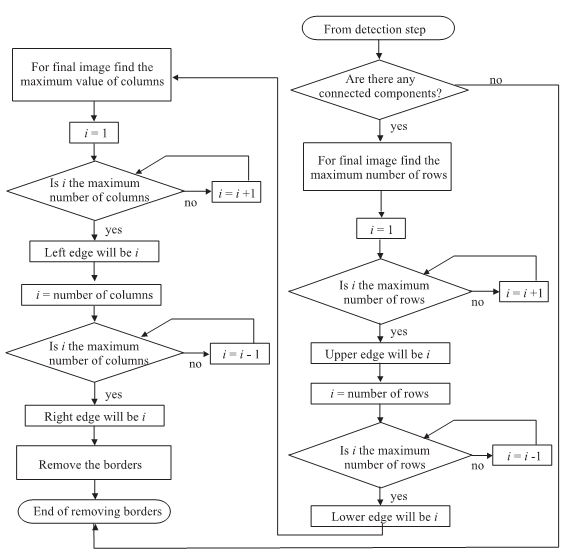


Figure 14: Flowchart for Removal of Border

**5.2** **Divide the text into rows**

After moving the border, the area has been divided into rows. Each row is saved in the next step figure.

Both above steps were done in a single nested loop for image. We have scanned the image vertically and horizontally in a single nested loop in this way we increases the processing speed of application and all the starting ending coordinates of row were stored in multidimensional array system we will describe it letter.

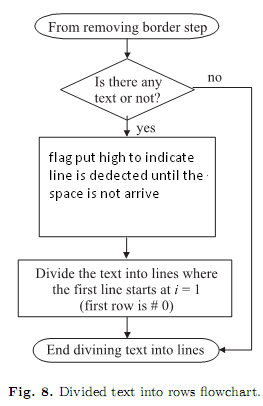


Figure 15: Flowchart for dividing Text in to Lines

**5.3 Divide the rows (lines) into words**

The single line is then divided into words. Before that**,** the empty area before and after the text are removed. Word and letters are distinguished behave of the spaces between the two letters. Again all the information of coordinates of word is stored in array again.

Flowchart for dividing rows into words is drawn below

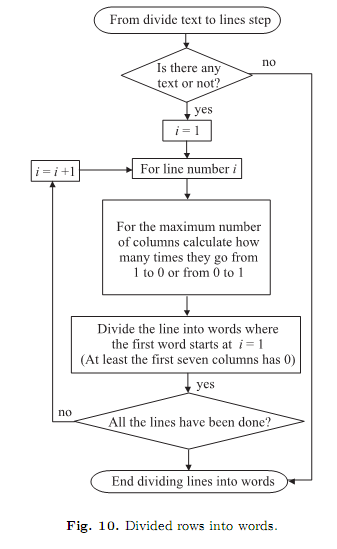


Figure 16: Flowchart for dividing line into words

**5.4 Divide the word into letters**

Each word is then divided into characters and saved in an array. This array is represents an image of latter. This image is again resized as we know that a text material can have different font size letters. So before we move this letter in reorganization system we resize each letter.

Flowchart for dividing words into characters is drawn below

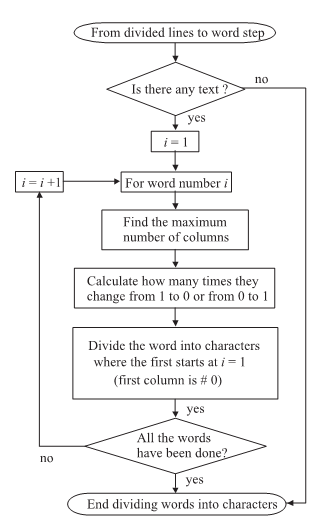


Figure 17: Flowchart for dividing word into Characters

As per coding aspect above both step again implemented in a single nested loop, here we excess each row(lines) from previous records, then each row is vertically scanned we distinguish the letters and words in each row and the process is repeated for all the rows. After finding the horizontal boundary of letters we recheck again for the vertical boundary as we know letters can be uppercase or lowercase in same line so the vertical height of different letters can vary in same line so we introduce concept of chopping in extracted letters.

**CHAPTER 4**

**PATTERN RECOGNITION IN OCR SYSTEM**

It is often useful to have a machine perform pattern recognition which is important in sense to apply to extracted images in extraction stages. In particular, machines that can read symbols are very cost effective. A machine that reads banking checks can process many more checks than a human being in the same time. This kind of application saves time and money, and eliminates the requirement that a human perform such a repetitive task

**6.1 Pattern recognition using correlation method**

We have applied correlation technique for Text recognition in our project. It finds the correlation of extracted character with that of every possible letter in database. Thus, finds the letter in database to which extracted letter correlation is maximum and prints that letter to text file which is further converted to speech using Window Api application in MATAB.

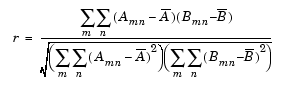
**6.1.1Description**

Correlation quantifies the strength of a linear relationship between two variables. When there is no correlation between the two quantities, then there is no tendency for the values of one quantity to increase or decrease with the values of the second quantity.

The main result of a correlation is called the **correlation coefficient** (or "r"). It ranges from -1.0 to +1.0. The closer r is to +1 or -1, the more closely the two variables are related. If r is close to 0, it means there is no relationship between the variables. If r is positive, it means that as one variable gets larger the other gets larger. If r is negative it means that as one gets larger, the other gets smaller (often called an "inverse" correlation).

In MATLAB r = corr2 (A, B) computes the correlation coefficient between A and B, where A and B are matrices or vectors of the same size.

Corr2 computes the correlation coefficient using



Where = mean2 (A), and = mean2 (B).

**6.2 Character Recognition using neural network**

A network is to be designed and trained to recognize the 26 letters of the alphabet. An imaging system that digitizes each letter centered in the system's field of vision is available. The result is that each letter is represented as a 5 by 7 grid of Boolean values. However, the imaging system is not perfect, and the letters can suffer from noise.

A network can be designed and trained to recognize the 26 letters of the alphabet. An imaging system that digitizes each letter centred in the system's field of vision is available. The result is that each letter is represented as a 5 by 7 grid of Boolean values. However, the imaging system is not perfect, and the letters can suffer from noise.

Perfect classification of ideal input vectors is required and reasonably accurate classification of noisy vectors. The twenty-six 35-element input vectors are defined as a matrix of input vectors called alphabet. The target vectors are also defined in this file with variable called targets. Each target vector is a 26-element vector with a 1 in the position of the letter it represents, and 0's everywhere else. For example, the letter A is to be represented by a 1 in the first element (as A is the first letter of the alphabet), and 0's in elements two through twenty-six.

6.2.1 Neural Network

The network receives the 35 Boolean values as a 35-element input vector. It is then required to identify the letter by responding with a 26-element output vector. The 26 elements of the output vector each represent a letter. To operate correctly, the network should respond with a 1 in the position of the letter being presented to the network. All other values in the output vector should be 0.

In addition, the network should be able to handle noise. In practice, the network does not receive a perfect Boolean vector as input. Specifically, the network should make as few mistakes as possible when classifying vectors with noise of mean 0 and standard deviation of 0.2 or less.

6.2.2 Architecture

The neural network needs 35 inputs and 26 neurons in its output layer to identify the letters. The network is a two-layer log-sigmoid/log-sigmoid network. The log-sigmoid transfer function was picked because its output range (0 to 1) is perfect for learning to output Boolean values.

The hidden (first) layer has 25 neurons. This number was picked by guesswork and experience. If the network has trouble learning, then neurons can be added to this layer. If the network solves the problem well, but a smaller more efficient network is desired, fewer neurons could be tried.

The network is trained to output a 1 in the correct position of the output vector and to fill the rest of the output vector with 0's. However, noisy input vectors can result in the network's not creating perfect 1's and 0's. After the network is trained the output is passed through the competitive transfer function compet. This makes sure that the output corresponding to the letter most like the noisy input vector takes on a value of 1, and all others have a value of 0. The result of this post processing is the output that is actually used.

6.2.3 Training

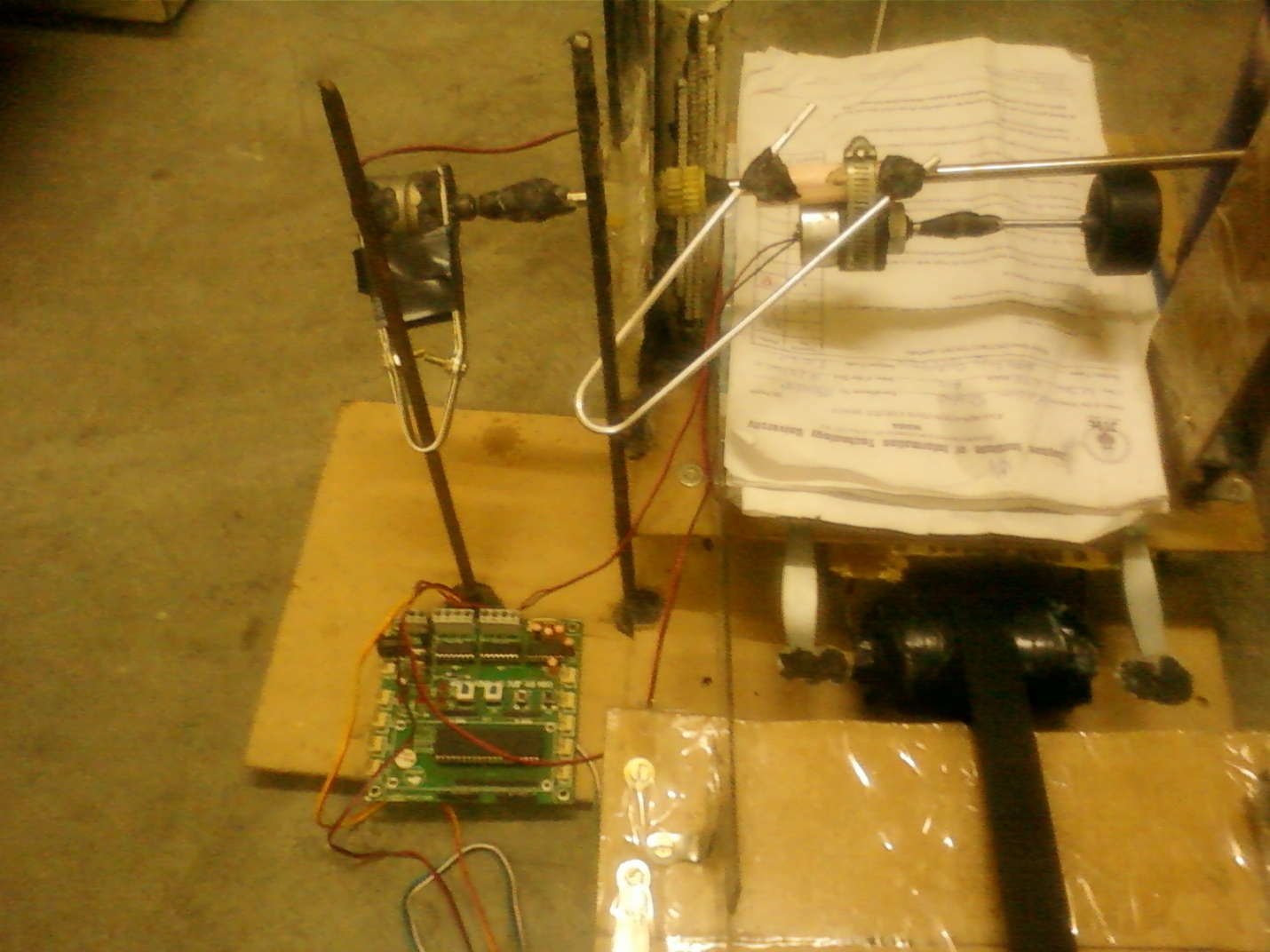
To create a network that can handle noisy input vectors, it is best to train the network on both ideal and noisy vectors. To do this, the network is first trained on ideal vectors until it has a low sum squared error. Training is very crucial since in this weights are assigned for different inputs.

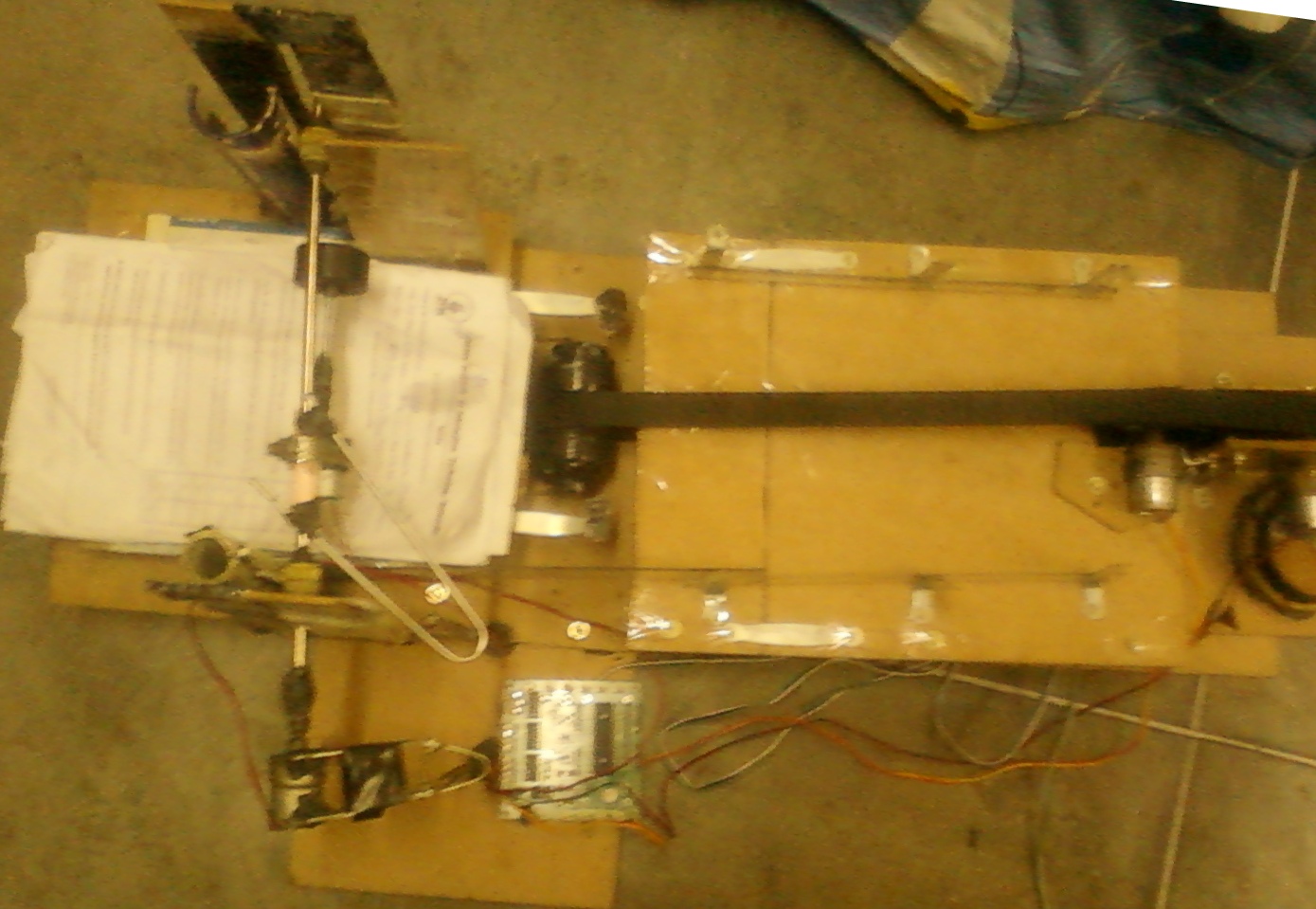
Then the network is trained on 10 sets of ideal and noisy vectors. The network is trained on two copies of the noise-free alphabet at the same time as it is trained on noisy vectors. The two copies of the noise-free alphabet are used to maintain the network's ability to classify ideal input vectors. All training is done using back propagation with both adaptive learning rate and momentum.

**CHAPTER 5**

**HARDWARE DESIGN**

In Hardware a system has been designed to collect a bundle of papers and process them one by one. Of the whole lot a single paper is selected and moved forward .The hard ware module has 4 motors connected to Microcontroller ATMEGA16 through the interface of motor drivers L293B. Microcontroller is further connected to computer by serial communication with MAX232 as interface. Of the Four motors attached the first motor is used for vertical motion, to attain vertical motion Rack & Pin mechanism was used, this vertical motion helps to place a bundle of papers in a tray. The second motor has a wheel attached to its axle which helps in selecting a single paper out of a whole lot. The paper that is separated from a lot is placed on conveyor belt which carries the paper from one end of the system to the other. The belt is connected to two motors which are in synchronisation with each other .The belt rotates in both clock wise and Anti-clock wise directions to have proper synchronisation with webcam,so that webcam can take pics to collect the data from the Answer sheets.



 A webcam is mounted above the belt which is connected to the computer but has a synchronisation with microcontroller through Serial Communication. The Images taken by the webcam are stored in the computer and further processed. The papers that are processed were moved forward and collected in a tray. Thus all papers placed in the tray were processed and collected at the end. To have an effective access to the device a voice module has been attached to give voice instructions to the users. Of all the voice modules available Apr9600 module has been selected as it has playback and record time of 60sec. This module can be interfaced with microcontroller which in turn is connected in computer through serial communication. The microcontroller controls the operations such as record and playback according the requirement. The data that is recorded can also be processed further, this provides user interface and convenience. To control the data flow in Apr9600 the ADC port of Atmega16 is configured as Output. To make the message ports active ‘0’ should be sent through ADC ports ,if the data overflow can be identified with a busy indication .

We can work in two modes

1.Random Access Mode

2.Tape Mode

**9.1 Random Access Mode**

In this mode recording or playback can be made randomly in any of the selected messages. The length of each message segment is the total recording length available (as defined by the selected sampling rate) divided by the total number of segments enabled.

9.1.1 Functional Description of Recording in Random Access Mode

On power up, the device is ready to record or play back, in any of the enabled message segments. To record, /CE must be set low to enable the device and /RE must be set low to enable recording. Recording can be initiated by applying a low level on the message trigger pin that represents the message segment we intend to use.

9.1.2 FunctionalDescription of Playback in Random Access Mode

To playback, /CE must be set low to enable the device and /RE must be set high to disable recording & enable playback.Playback can be initiated by applying a high to low edge on the message trigger pin that representing the message segment you intend to playback. Playback will continue until the end of the message is reached. If a high to low edge occurs on the same message trigger pin during playback, playback of the current message stops immediately.

**9.2 Tape Mode**

Tape mode manages messages sequentially much like traditional cassette tape recorders. Within tape mode two options exist, auto rewind and normal. Auto rewind mode configures the device to automatically rewind to the beginning of the message immediately following recording or playback of the message. In tape mode, using either option, messages must be recorded or played back sequentially, much like a traditional cassette tape recorder.

9.2.1 Function Description Recording in Tape Mode

On power up, the device is ready to record or play back, starting at the first address in the memory array. To record, /CE must be set low to enable the device and /RE must be set low to enable recording. A falling edge of the /M1\_Message pin initiates voice recording (indicated by one beep). A subsequent rising edge of the /M1\_Message pin during recording stops the recording (also indicated by one beep). If the /M1\_Message pin is held low beyond the end of the available memory.

9.2.2 Function Description of Playback in Tape Mode

On power-up, the device is ready to record or play back, starting at the first address in the memory array. Before we begin playback, the /CE input must be set to low to enable the device and /RE must be set to high to disable recording and enable playback. The first high to low going pulse of the /M1\_Message pin initiates playback from the beginning of the current message; on power up the first message is the current message. When the /M1\_Message pin pulses low the second time, playback of the current message stops immediately.

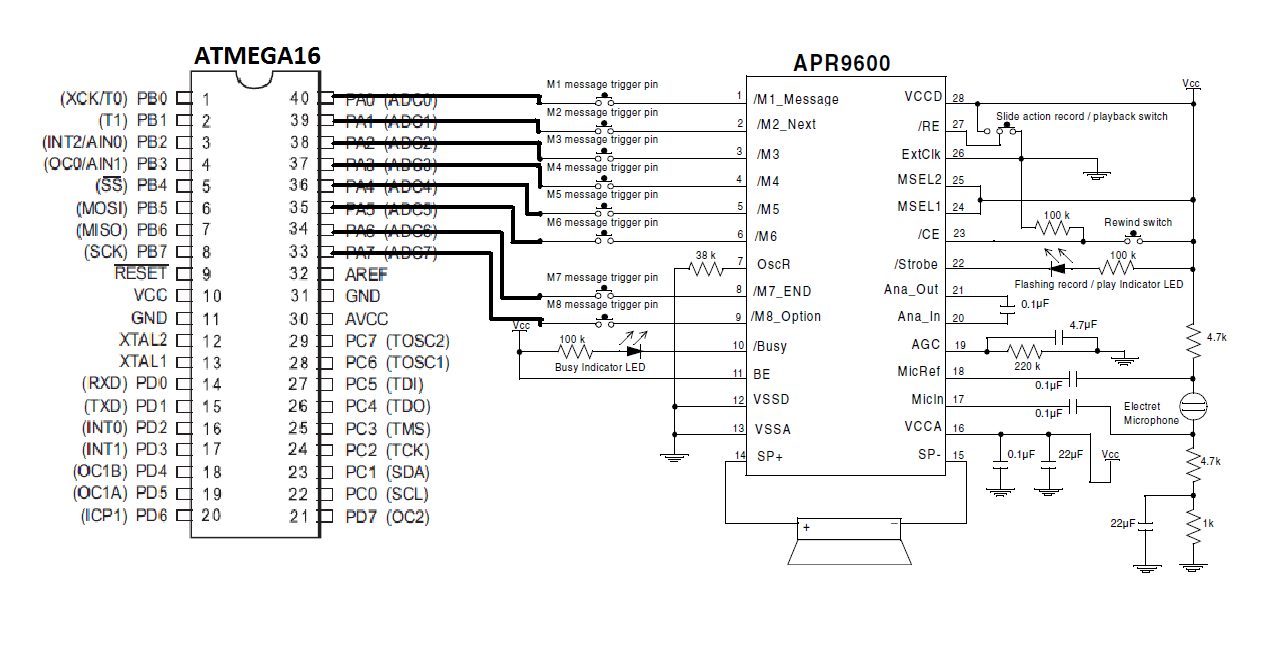


Figure 20: Interfacing of APR9600 with ATMEGA16

|  |  |  |  |
| --- | --- | --- | --- |
| PIN NO | PIN NAME | **Normal Option** | **Auto Rewind Option** |
| 1,2,3,4  5,6 | Message1,2,  3,4,5,6 | Message: A low edge on this pin plays or records the next message. | Message: A low edge on this pin  plays or records the current message. |
| 8 | M7\_END | During playback a low level on this pin indicates that all recorded messages have been played. During recording a low level on this pin indicates  that the end of the memory  array was reached. | During playback a low level on this  pin indicates that all recorded messages  have been played. During recording a low level on this pin indicates that the end of the memory array was reached. |
| 9 | M8\_Option | Option: This pin in conjunction with MSEL1 and MSEL2 sets record and playback operating mode. Consult table 1 for decoding information. | MSEL1 and MSEL2 sets record and  playback operating mode.  . |

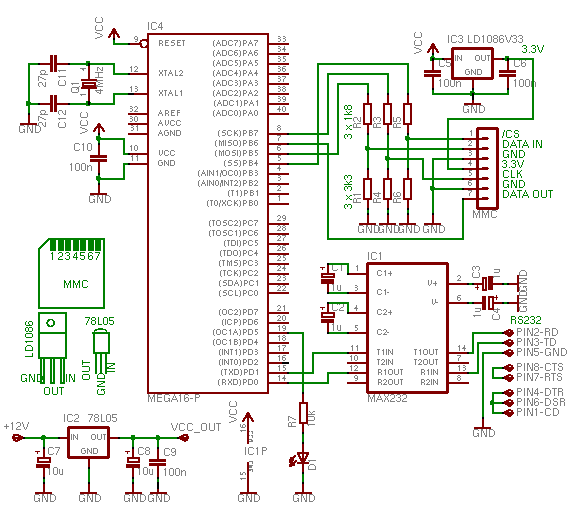


Figure 21: Serial Interface between ATMEGA16 and Computer

Serial Port is a serial communication interface through which information transfers in or out one bit at a time(Contrast parallel port). Throughout the history of personal computers, data transfer through serial ports connected the computer to devides such as terminals and various peripherals. The serial port usually identifies hardware more or less compliant to RS-232 standard, intended to interface with a communication device.

ATmega16 provides three subsystems for serial communications.

1. Universal Synchronous & Asynchronous Serial Receiver & Transmitter (USART)

2. Serial Peripheral Interface (SPI)

3. Two-wire Serial Interface (TWI)

USART :

It supports full-duplex mode between a receiver and transmitter and is typically used in asynchronous communication. Start bit and stop bit are used for each byte of data.

Max 232 IC is selected as an interface between computer and ATmega16. Max232 has a capacitive voltage generator to supply TIA/EIA-232-F voltage levels from a single 5-v supply.

USART of ATmega16 supports baud rates from 960bps up to 57.6kbps and a character size of 5 to 9 bits, 1start bit, 1or 2 stop bits and a parity bit which is optional.

Steps Involved in synchronizing serial port:

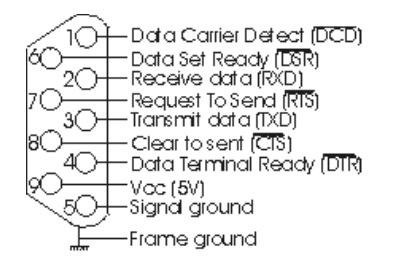
1.Intializing the serial port.

* Set USART communication parameter (data bit, stop bit, parity bit)
* Enable transmitter and receiver
* Set USART for asynchronous mode.
* Set Baud rate

2.Sending a character.

3.Receiving a character.

4.Sending / receiving formatted strings.

**MOTOR-3**

**MOTOR-1**

**MOTOR-4**

**MOTOR-2**

**Matlab**

**Environment**

**Micro Controller**

**(ATMEGA-16)**

**IR SENSOR**

**CAMERA**

**CHAPTER 6**

**CHALLENGES FACED AND SOLUTIONS SUGGESTED**

1. Bubbles may exist inside the extracted characters due to scratches, reflection or nails on the picture taken by camera. In these situations, a bubble filling algorithm can be used to fill the bubbles. Therefore, the filling algorithm can be used based on both labelling and area calculation. However, the labelling algorithm will pose additional computational burden with no considerable effect on recognition. Thus, this step has not been adopted, in our project, in order to reduce the processing time.

2. There can be large variation in both the font size and font type of the text expected in the diverse forms of images captured. Therefore the threshold box for segmentation cannot be fixed at a specific size.

3. Resolution of such images will be typically modest. Coupled with an uncontrolled environment, uneven illumination and reflection and a possibly odd image capturing angle, the target text captured can be blurred, all of these posing difficulties to text extraction.

4. The text extraction and recognition function will inevitably be limited by the nature of the small-screen mobile devices, which will restrict the span of the image which can be captured. Finally, images taken under a poor lighting condition may result in low entropy. Low entropy may also cause problems in processing. Far more intensive computation will be necessary when the text is embedded in a complex background

5. Calculating the histogram, and then smooth it to ﬁnd the threshold points. These methods are not working well when the text and background colours are very similar, and also if it may create an area of negative text. Therefore histogram technique is not implemented rather mask defined in pre-processing stage is implemented.

6. **No support for handwritten text** OCR has severe limitations when come to handwritten text. Characters must be handwritten with separate characters in different box.

7.We found better solution to existing solutions when connected components are considered .If length and width ratio of extracted character is greater to that of critical ratio then image can be divided in to partial images, the way of dividing connected characters and treating them as partial images has increased the accuracy and time consumption has been decreased.

8.Implementation of Erosion and Skeletonisation technique together for the identification of characters has solved the problem of fake connections of two different characters to some extent.

**CHAPTER 7**

**FUTURE SCOPE**

OCR is one of the most emerging technology and its reliability is continually improving. Soon OCR will become a powerful tool for data entry applications which will lead to automated data entry by OCR,thus reducing labor.Incorporating OCR will be an attractive feature of any Data Entry System.However in past due to limited availability of a capital and short environment was restricting the growth of this technology,but today more and more enterprises are working on this technology and that will definitely lead to 100% accuracy in this technology thus making the dream of paperless world true.

Our work has been defined to only fewer fonts at present,In future we would extract the text from hand written images and also from all the other fonts.Templates were being designed to make the task easier.TTS accuracy can also be increased to such extent that it could become user interactive.

Following OCR algorithm can be used for development of a human-machine interactive software application, specifically useful for text extraction from images, which are captured using mobile and digital devices with cameras which can be used to translate the text in such a captured image to another language.

1. The approach is also viable as an alternative way to send SMS from images captured and to extract URL text from a complex background and directly link the user to the website via his mobile device.

2. Implementation of **spell checking** technique can greatly enhance the working of OCR system by increasing its recognition accuracy.

3. Increasing the efficiency or enhancing the box called creating zone will also increase the efficiency of software.

4. Colour detection will also be useful to help user to reduce job of formatting in any sort of text.

**Applications of OCR**

* Avoids Re-Typing of a Page
* Format a paper that was Faxed.
* Enter an article into bibliographic database.
* Can make the softcopies for all the documents easily.
* Preservation of old manuscript.
* Computer meter reading.

**REFERENCES**

1. Text Detection and Character Recognition using Fuzzy image processing by Mohanad Alata — Mohammad Al-Shabi, Journal of ELECTRICAL ENGINEERING, VOL. 57, NO. 5, 2006, 258–267.

2) YUAN, Q.—TAN, C. : Text Extraction from Gray Scale Document Images Using Edge Information, Proc. Sixth Int. Conf. on Document Analysis and Recognition, 302–306.

3) WOLF,C.—JOLION,J.—CHASSAING,F. : Text Localization, Enhancement and Binarization in Multimedia Documents,In Proceedings of the International Conference on Pattern Recognition (ICPR). IEEE Computer Society, Canada: Quebec City. 2002, 4, 1037–1040.

4) Canny, J. (1995) ‘A computational approach to edge detection’, *IEEE Trans. on Pattern Recognition*, Vol. 17, No. 12.

5) Gonzalez, R.C. and Woods, R.E. (2001) *Digital Image Processing*, 2nd ed., Chapter 6, Chapter 10,Prentice-Hall, Englewood Cliffs, NJ.

6) Y. Deng and B. S. Manjunath. Unsupervised segmentation of color-texture regions in images and video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*,23(8):800{810, 2001.

7) Y. Mori, H. Takahashi, and R. Oka. Image-to-word transformation based on dividing and vector quantizing images with words. In *Proc. of First International Workshop on Multimedia Intelligent Storage and Retrieval Management*, 1999.

8) Arth, C., Limberger, F. and Bischof, H., 2007. Real-time license plate recognition on an embedded DSP-platform. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR ’07) pp. 1–8..

9) Wolf, C., michel Jolion, J. and Chassaing, F., 2002. Text localization,enhancement and binarization in multimedia documents.In Proceedings of the International Conference on Pattern Recognition (ICPR) 2002, pp. 1037–1040.

10) Anderson, J. A., 1995, *Introduction to Neural Networks* (Cambridge, MA:MIT Press).

11) Using Neural Networks to Create an Adaptive Character Recognition System Alexander J. Faaborg Cornell University, Ithaca NY.