AIVP

ADVANCES IN IMAGE AND VIDEO PROCESSING

TABLE OF CONTENTS

DISCLAIMER

All the contributions are published in good faith and intentions to promote and encourage research activities around the globe. The contributions are property of their respective authors/owners and the journal is not responsible for any content that hurts someone's views or feelings etc.

A Camera System for Detecting Dust and Other Deposits on Solar Panels

E. A. Yfantis[1] and A. Fayed[2]

[1] *Computer Science Department, University of Nevada, Las Vegas, Nevada, USA*
[2] *Mechanical Engineering Department, University of Nevada, Las Vegas, Nevada, USA*
yfantis@cs.unlv.edu

ABSTRACT

Solar panels over time, due to winds, sandstorms, bird droppings, suffer from dust, and other deposits. As a result of these deposits the sunlight is refracted, and only part of the sunlight reaches the chips inside the glass cage that are generating the electricity. In a remote area with thousands of solar panels, it is both expensive and cumbersome to send maintenance people to inspect each panel and clean it if needed. We have smart cameras with R, G, B, and infrared for night vision, that take the picture of each panel continuously. The picture becomes input to our classification algorithm that decides real time if the panel needs cleaning or not. Our classification algorithm consists of: our classification vector, the metric used, the training of the classifier, the testing of the classifier, and the classifier put into play for everyday use. At the present time we use a commercial camera transmitting JPEG frames wireless to our server where the classification and storage takes place. But in the near future our classification algorithm will reside on a flash memory which will be part of a circuit board that we are designing. The algorithm operates on the incoming data and will be executed by an ARM processor which will also be on the board. The circuit board also will include a CCD and Infrared camera. The hardware and software on this electronic board will be designed and programmed by the authors. Once our intelligent system detects that the panel needs cleaning it will automatically trigger a mechanism which will clean the panel.

Keywords—Multivariate distribution; Mahalanobis distance; Hotelling's T-square; misclassification.

1   Introduction

In a large scale solar energy array of panels build by a variety of solar panel manufacturers, each panel having a different area, and based on different type of chip architectures, the energy output is not uniform and there could be a considerable panel to panel variation. The energy output of each panel depends on the time of the day, day of the month, month of the year, year in the solar cycle, and of course the climatological conditions during the time of the day that energy measurements are obtained. In arid climates it rains very rarely so the ground is dry therefore when the wind blows the loose ground is carried by the wind and dust storms are very common. As the dust settles on the glass of the various panels the energy output falls because the dust prevents part of the light incident to the glass to enter and reach the chips generating the energy. Thus part of the light is refracted and less light goes through the glass and into chip structure designed to convert the solar energy to electricity. The loss of light energy depends on the amount of dust, the size of particulates, and the chemical composition of dust. Loss of solar energy is not just due to dust but also due to bird droppings and other factors that are yet to be explained. In arid climates there are not as many trees

as in climates with normal amount of rain. As the birds migrate from north to warmer climates during the fall, or from warm to cooler climates during the spring, solar farms are used by the migrating birds as rest areas, and therefore get many bird dropping that reduce the electric energy output. Also during the regular season solar panels get a fair amount of bird droppings. In arid climates that often times the summers are too hot, and the winters nights are cold, solar panels provide shade in the summer and protection from the cold and the wind in the winter. Thus many small animals find rescue from the difficult climate conditions and make home under the panels. Often times some of these animals try to eat the cables attached to the panels and in the process damage the cable to the point that stops the flow of electricity from the panel to the destination. In addition to the animals and birds, the protections the panels provide from the harsh weather conditions promote the growth of various desert plants that grow faster in the shaded ground underneath the panels. Often times these plants and trees grow enough to crack the base of the panel, or lift the panel so that the orientation of the panel with respect to the sun is not what it should be, therefore   the amount of light going through the glass and into the chip structure designed to convert the solar energy to electricity is less than the optimal amount. Often times these plants as they grow into trees they lift the panel and severely damage it as well as other neighboring panels that this panel falls on the top of. In addition to the above possibilities for damaging the panels and reducing the electric energy output there are others, such as vandalism, earthquakes, theft, and a number of catastrophes considered to be acts of God. For now the only purpose of our smart camera is to detect if a panel needs cleaning by applying our classifier and then if it does need cleaning to trigger a mechanism to clean the panel. As we looked at a number of cleaning mechanisms including a container with water and two sponges one to wash the panel with the water and the other to dry, or a container with a cleaning fluid and two sponges, or a mechanism consisting of small spike-like needles to crate static electricity on the dirt of the panel and attract the dirt into a vacuum, we concluded that the later is the best method and has many advantages over the previous mentioned methods. Some of the advantages of the static electricity mechanism with the vacuum are its very low maintenance. The mechanism with the water and the sponges need enormous amount of maintenance in an arid climate. The sponges dry out relatively fast in the extreme climate conditions of the summer months that the temperature could rise well over 110 degrees Fahrenheit and could need to be replaced very often adding to the expense of maintenance of the solar panel site. Furthermore water is very scarce, and evaporates very fast as it is poured on the top of the panels. As the mechanism cleans the panel with water is very possible that it would create uneven cleaning and/or leave a film of material that acts is inhibitors for the panel to utilize the solar energy in an optimal way. Replacing the water in the containers of each solar panel could be expensive and challenging. It might have an initial expense to bring the water pipe into the site and then a continuous expense to maintain the containers and their sensors that allow the proper release of the water. If a cleaning fluid is used the challenges are similar to those of using the water, but the expense could be much higher, and that includes cleaning fluid replacement, sponge and/or wiper replacement, and man power needed for the proper maintenance. The static electricity mechanism to remove the dust is very light, gets a small amount of energy from the panel itself , is relatively simple to build, therefore could be very reliable, and needed very little maintenance. In this research paper we are focusing only on building a classifier to detect if the panel needs to be cleaned or not. The steps for building our classifier included sampling, deciding which parameters our classification vector should include, train the classifier, compute the probability of misclassification which includes the probability of the classifier deciding that the panel is dirty

although it is not, or the panel is clean although it is not. After testing the classifier we put it into play on a real solar panel site. For now the classifier is computer software on the server where the camera transmits the video of the solar panels. In the near future the classifier will be embedded into a hardware board, which will include the visual camera chip, the infrared camera, and the alarm events that will be transmitted wirelessly and alarming the care takers of the site about broken glass, a malfunction in one of the panels, cut cables, or vandalisms. Our classifier is based on the multivariate probability distribution function of the mode of the red channel, the mode of the green channel, and the mode of the blue channel. It also includes the marginal distribution function of the modes of the red, green, and blue channels. In clean panels the three dimensional vector of the mean vector of the modes has relatively low values, while as the panel gets dirty by dust particulates and bird droppings, the means of the mode vectors of the red, green, and blue channels increase. A Hotelling's T-square test shows that the mean vectors of the modes between clean and dirty panels are significantly different. Also the Mahalanobis distance is capable of classifying correctly with high probability that a panel belongs to the clean or dirty category of panels. This paper is structured as follows: First is the abstract, section I is the Introduction, section II is the Background Information, Section III is the Classification Algorithm, Section IV is the Conclusions, Section V is the References.

2   Background Information

1.1   The color of the Sky

The sky color variations from shades of blue to shades of gray indicate higher levels of concentration of particulates larger than the clean air. The clean air is composed approximately of 78.084% Nitrogen, 20.94% Oxygen, 0.934% Argon, and 0.0350% other gases [1], [2]. The color of the sky is explained by the Mie solution to Maxwell's equation, and Rayleigh's scatter approximation theory. According to Rayleigh's scatter theory the small molecule size of Nitrogen and Oxygen molecules scatter the components of the light with shorter wavelength range and thus higher frequency. These light components in the visible spectrum are the blue and violet colors. The scattering of the blue colors therefore by Nitrogen and Oxygen explains the blue color of the sky. Larger molecules scatter color components with longer wavelength. In the visible light the red color component of the white light has the larger wavelength, then the green color component has lower wavelength than the red, and the blue color component has lower wavelength than the green. When the concentration of larger molecules such as pollutants increase, the scattering of colors with longer wavelength increases. The mix of the longer wavelength color components in the atmosphere with the blue, change the color of the sky, making it more grey, or light grey. Thus when there is high concentration of water vapor molecules that cling together in that part of the sky make it look white or gray. This is the reason that clouds look white or gray. When pollution levels increase, since their molecule size is usually relatively large, that triggers the scattering of higher wavelength color components. The color of the sky is a good indicator of how clean the air is. Lighter grey colors signify the presence of water vapor, sulfur, aerosols, and soot in the atmosphere. In any given direction skylight is light scattered by all the molecules and particles in its path from the sun which is the source of daylight in the Earth. Since the atmosphere on the Earth is composed mostly of Nitrogen and Oxygen molecules, then these are the most frequent molecules that light will encounter in its trajectory. Electromagnetic waves with shorter wavelengths are affected by particles much smaller than their wavelength, scattering these wave in different directions. Waves with larger wavelength are not being affected by these small particles. As a result the blue and violet waves of the white light are scattered while traveling in the clean atmosphere turning the sky blue [3],[4],[5]. The lower intensity of the violet color compared to other colors in the light coming from the sun, combined with the

lower sensitivity of this color by the human eyes is the main reason why the sky is blue and not violet even though the violet wavelength is smaller than the blue and scatters even more. So there is very little scattering of the red color due to its relatively large wavelength. The red color travels straight thru the atmosphere. Similarly there is very little scattering for the green color because although its wavelength is smaller than that of red is still relatively large compared to the size of Nitrogen and Oxygen. So the green light travels straight thru a clean atmosphere.

1.2   Light and Electric Energy

The infrared and red wavelengths having the larger wavelengths compared to the other components of the white light incident to the glass of the solar panel penetrates the glass and produces the most amount of electricity. The blue and violet wavelengths suffer the most absorption by the glass of the panel and contribute very little in the production of electric energy. As we pointed out in our previous section when the sky is cloudy due to water vapor, the red and green components of the light suffer a great deal of scattering and do not go straight thru the atmosphere. The infrared wavelengths are the ones that produce more energy following by the red, green, and blue. The energy produced by the red and green wavelengths is substantially reduced during cloudy days. Also when dust particulates, or other pollutants are in the atmosphere they cause the scattering of the larger wavelength components of light thus the energy produced by the solar panels is reduced. Dust storms are very common in arid climates, especially during the summer months when the ground is very dry and the top soil is sandy. Sandstorms are responsible for depositing dust and sand on the solar panels. As the sunlight passes thru these particulates the larger wavelengths that normally are the larger energy producers, scatter, and as a result of this scattering less light energy is reaching the chips producing the energy and therefore less energy is produced. Birds flying over the solar panel area, or stopping on solar panels for a rest or out of curiosity, often times they deposit their droppings. Bird droppings usually form a thick coating on an area of the panel preventing the sun light in that area from going thru. It is very difficult for one to change the weather or the pollution in the air. It is not difficult to detect the dust and sand on the glass and clean it. As we mentioned before, our approach is not to use water or other liquid for cleaning, but to use a small vacuum cleaner flying over the panel in a methodical way that goes thru in a raster type of way and visits every point of the panel from top to button twice. During the first pass it vacuums all the gravel, dust, and droppings it can. During The second pass it activates the needles that have static current, thus charging the dust and droppings on the glass with opposite static current. Thus our system visits every point on the glass in a raster scanning fashion, the needles attract the remaining deposits on the glass, and those deposits go thru the vacuum and get deposited thru a pipe system on the ground by the base of panel. This is a relatively simple system, draws the energy from the panel itself, does a very good job in cleaning, does not leave any film on the glass, has a relatively small manufacturing and installation cost, it needs very little maintenance. It is very easy to train a technician to repair or replace the system.

3   Classification Algorithm

As we examine an image of a clean panel (Fig. 1), and compute the probability density functions of the red channel, the green channel, and the blue channel of the pixels of the image (Fig. 3) we observe that each one of the marginal probability density functions for each one of the channels is skewed to the left with a maximum (mode) to the left of the 127.5 value. Here we assume that the scale of the red channel of the pixels is between 0-255, similarly the scales of the green and blue channels of the pixels are also between 0-255. Thus in the graphs of the probability density function

of the color channels the scale on the x-axis is between 0-255. Fig. 3(a) gives the univariate probability distribution functions of the red, green, and blue, channels, for a clean panel. The mode for the red channel is 56, for the green is 60, and for the blue is 84. The x-axis for each one of the channels is from 0-255. Fig. 3(b) gives the univariate probability distribution functions of the red, green, and blue channels, for a dusted panel. The mode for the red channel is 79, for the green is 82, and for the blue is 105. Based on a sample of 50 pictures of dirty panels and fifty of clean ones, the gamma distribution provides enough flexibility to model these probability functions. The gamma density function is of the form [6], [7]:

$$f(x) = \frac{x^{\alpha-1}e^{-\frac{x}{\beta}}}{\Gamma(\alpha)\beta^{\alpha}}, \ x > 0, \ \alpha, \beta > 0 \tag{III.1}$$

Theorem 1, gives an estimate of the global mode of the gamma density function in terms of its parameters α, β, and also estimates the parameters from the data.

Theorem 1. The global mode of the gamma density function (the point x that maximizes the density function f(x) is $x_m = (\alpha - 1)\beta$, if $\hat{x}$ is the global mode based on the empirical density function computed from the data, and $\bar{x}$ is the average then an estimate of the parameters α, β, from the data is: $\hat{\alpha} = \frac{\bar{x}}{\bar{x}-\hat{x}}$ and $\hat{\beta} = \bar{x} - \hat{x}$

Proof

The global mode x_m of the gamma probability function is the point that maximizes the density function. Namely the point for which the first derivative of the function f(x) is zero and the second derivative is negative.

$$f'(x) = e^{-\frac{x}{\beta}} \frac{(\alpha-1)x^{\alpha-2} - \frac{x^{\alpha-1}}{\beta}}{\Gamma(\alpha)\beta^{\alpha}} = 0 \tag{III.2}$$

The solution to the above equation is

$$x_m = (\alpha - 1)\beta \tag{III.3}$$

$$f''(x) = -\frac{1}{\beta}e^{-\frac{x}{\beta}} \frac{(\alpha-1)x^{\alpha-2} - \frac{x^{\alpha-1}}{\beta}}{\Gamma(\alpha)\beta^{\alpha}} + e^{-\frac{x}{\beta}} \frac{(\alpha-1)(\alpha-2)x^{\alpha-3} - \frac{(\alpha-1)x^{\alpha-2}}{\beta}}{\Gamma(\alpha)\beta^{\alpha}} \tag{III.4}$$

For the type of gamma probability function that fits this data α > 2 always. Thus

$$f''(x_m) = -e^{-(\alpha-1)} \frac{((\alpha-1)\beta)^{\alpha-3}}{\Gamma(\alpha)\beta^{\alpha}} < 0 \tag{III.5}$$

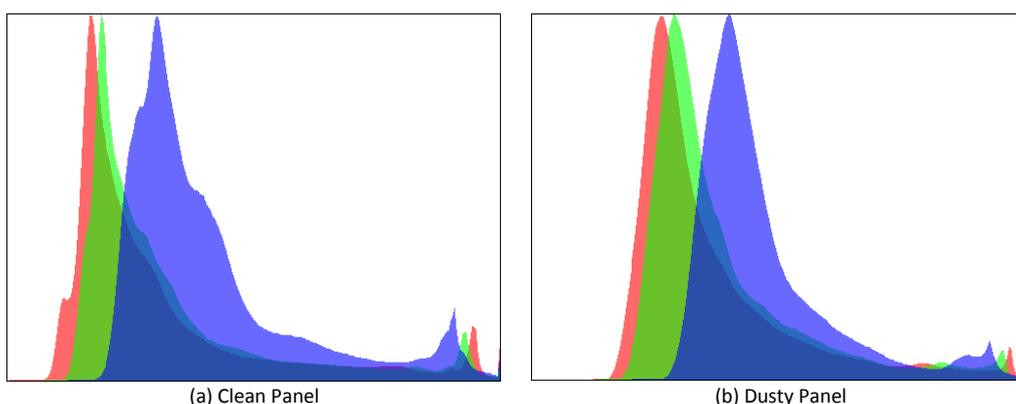Figure 1: Image of a clean panel   Figure 2: Image of a dusty Panel

(a) Clean Panel   (b) Dusty Panel

Figure 3: Probability density function for the R, G, B channels, Observe the shift of the red, green, and blue modes.

(a) Clean, Red   (b) Dusty, Red

(c) Clean, Green   (d) Dusty, Green

(e) Clean, Blue   (f) Dusty, Blue
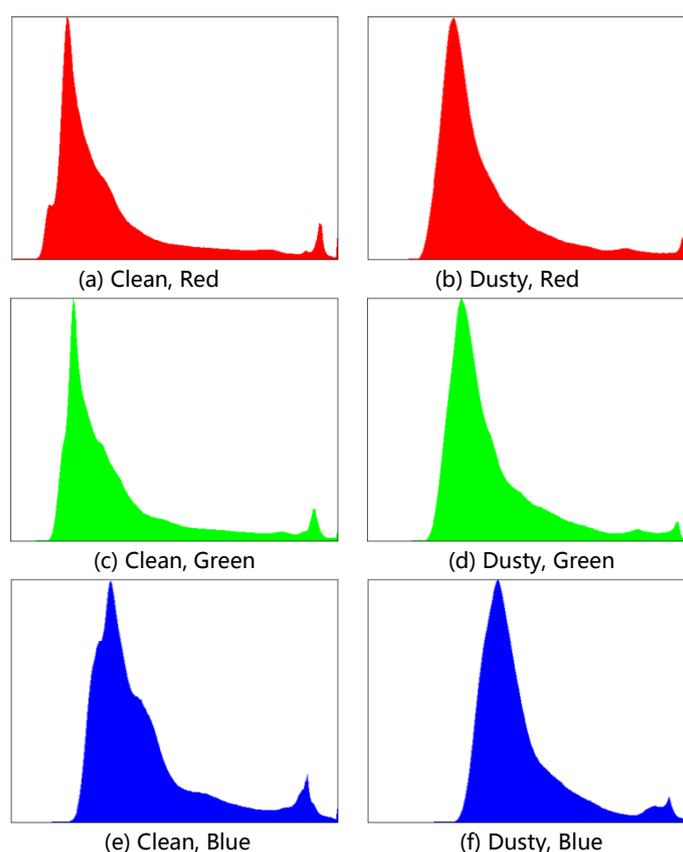
Figure 4: Probability density function for the R, G, B channels

The point $x_m = (\alpha - 1)\beta$ is the global maximum for f(x), thus it is the mode of the density function. It is very easy to prove that the mean or expected value of a gamma is μ=αβ . Given a panel image we can use the pixels to extract the values of the red, green, and blue channels. For each one of these channels we can estimate the mean, and the maximum mode. The mean is the average of all the values for that channel, and the maximum mode of course is the point between 0 and 255 for which most of the values of that channel for the picture that the channel belongs to, are equal to. Let the sample mean for the arbitrary channel be denoted by $\bar{x}$ and let the estimate of the maximum mode be denoted by $\hat{x}$. Then the parameters α, β, of the gamma can be estimated by setting:

$$\bar{x} = \alpha\beta \qquad\qquad\qquad (III.6)$$

and

$$\hat{x} = (\alpha\text{-}1)\beta \qquad \text{(III.7)}$$

From the above two equations we obtain

$$\bar{x} - \beta = \hat{x} \qquad \text{(III.8)}$$

which implies that:

$$\hat{\beta} = \bar{x} - \hat{x} \qquad \text{(III.9)}$$

From equation III.6 therefore we obtain that

$$\hat{\alpha} = \frac{\bar{x}}{\bar{x}-\hat{x}} \qquad \text{(III.10)}$$

QED.

Formulas III.9 and III.10 give us estimates of the parameters α and β which they define the gamma density function for the channel based on the channel data. We took the picture of each one of the fifty panels when they were dusty, then we cleaned them and we took their picture again. Thus we obtained 50 clean panel pictures and fifty dusty under the same conditions. For each one of the pictures we computed the probability density function of the red, green, and blue, channels. As we graphed each one of the probability density functions the x-axis denotes all the possible values a color-channel can have. Thus the values for each one of the three channels vary between 0-255. The value in the x-axis for which the experimental density function obtains its global maximum is the mode. For the clean pictures we obtained 50 mode data for the red channel, 50 mode data for the green, and 50 mode data for the blue. In each one of the pictures the mode for the blue is over 20 units higher than the mode for the green and the mode for the green is about 4 units higher than the mode for red. As the panels get dusty then the modes are shifted to higher values, and the electric energy of the panel is reduced. There is a tradeoff between the energy needed for the vacuum to clean the panel and the energy gained. There is a cutoff point beyond which not cleaning the panel significant energy is lost and therefore we need to detect that the panel has reached that cutoff point and trigger the mechanism automatically to clean the panel. The classification theory that enables us to make this decision is based on the following statistics. Consider the 50 clean panels. Let $(r_{m1}, r_{m2}, \dots, r_{m50})$, $(g_{m1}, g_{m2}, \dots, g_{m50})$ and $(b_{m1}, b_{m2}, \dots, b_{m50})$, be the 50 computed mode values for each one of the clean channels. Let $\bar{r}_m, \bar{g}_m, \bar{b}_m$, denote the average of the red-modes, green-modes, and blue-modes respectively. Then according to the central limit theorem the probability distribution function, of the vector $\bar{r}_m, \bar{g}_m, \bar{b}_m$ approaches to a trivariate normal. The random vectors $(r_{m1}, g_{m1}, b_{m1}), (r_{m2}, g_{m2}, b_{m2}), \dots, (r_{m50}, g_{m50}, b_{m50})$, are independent since each panel is autonomous and independent of the other panels, and identically distributed, with a mean vector $\underline{\mu}' = (\mu_{rm}, \mu_{gm}, \mu_{bm})$, where the $\underline{\mu}'$ signifies the transpose vector, and variance- covariance matrix

$$\Sigma = \begin{bmatrix} \sigma^2_{rm} & \sigma_{rmgm} & \sigma_{rmbm} \\ \sigma_{rmgm} & \sigma^2_{gm} & \sigma_{gmbm} \\ \sigma_{rmbm} & \sigma_{gmbm} & \sigma^2_{bm} \end{bmatrix} \qquad \text{(III.11)}$$

The expected value $E(\bar{r}_m, \bar{g}_m, \bar{b}_m) = \underline{\mu}' = (\mu_{rm}, \mu_{gm}, \mu_{bm})$, and the variance covariance matrix of the vector $\bar{r}_m, \bar{g}_m, \bar{b}_m$ is $\frac{\Sigma}{50}$. In general if n-pictures are used to compute the vector $\bar{r}_m, \bar{g}_m, \bar{b}_m$, then the variance-covariance vector of $\bar{r}_m, \bar{g}_m, \bar{b}_m$, would be $\frac{\Sigma}{n}$. The vector $\underline{\mu}' =$

$(\mu_{rm}, \mu_{gm}, \mu_{bm})$ is estimated from the data. Thus if a sample of n clean panels is used and we denote the trivare vector of the mode of the ith panel as $x_{mi} = (r_{mi}, g_{mi}, b_{mi})$ then

$$\bar{x}_m = \frac{x_{m1} + x_{m2} \ldots + x_{mn}}{n} \tag{III.12}$$

And the variance covariance matrix Σ can be estimated from the data as follows:

$$\hat{\Sigma} = \frac{1}{n-1} \sum_{i=1}^{n} (x_{mi} - \bar{x}_m)(x_{mi} - \bar{x}_m)' \tag{III.13}$$

Theorem 2 creates the criterion for classifying a panel as being clean or needing cleaning.

Theorem 2. Let x denote the vector of r,g,b, modes of a new panel and $t^2 = \frac{n}{n+1}(x - \bar{x}_m)'\hat{\Sigma}^{-1}(x - \bar{x}_m)$ then $\frac{n-3}{3(n-1)}t^2 \sim F_{3,n-3}$ is an F distribution with parameters 3, n-3. If $x'\hat{\Sigma}^{-1}x < \bar{x}_m'\hat{\Sigma}^{-1}\bar{x}_m + \frac{3(n^2-1)}{n(n-3)}F_{3,n-3,\varsigma}$, then the panel is classified as clean, otherwise as needed to be cleaned. $F_{3,n-3,\varsigma}$ in the classifier formula is the F-distribution value with parameters 3 and n-3, and level of significance ς.

Proof

Since x is the vector of r, g, b modes of a new panel, is independent of the n clean panels already used to computer $\bar{x}_m$. The

$$E(x - \bar{x}_m) = E\big((x - \mu) - (\bar{x}_m - \mu)\big) = E(x - \mu) - E(\bar{x}_m - \mu) = 0 \tag{III.14}$$

 and the variance covariance matrix of the vector $x - \bar{x}_m$ is

$E(x - \bar{x}_m)(x - \bar{x}_m)' = E[(x - \mu) - (\bar{x}_m - \mu)][(x - \mu) - (\bar{x}_m - \mu)]' = E(x - \mu)(x - \mu)' + E(\bar{x}_m - \mu)(\bar{x}_m - \mu)' - E(x - \mu)(\bar{x}_m - \mu)' - E(x - \mu)'(\bar{x}_m - \mu) = \Sigma + \frac{1}{n}\Sigma - 0 - 0 = (1 + \frac{1}{n})\Sigma$

$$\tag{III.15}$$

Notice that

$$(x - \mu), \ and \ (\bar{x}_m - \mu)'$$

are statistically independent therefore

$$E(x - \mu)(\bar{x}_m - \mu)' = 0 \tag{III.16}$$

For the same reason

$$E(x - \mu)'(\bar{x}_m - \mu) = 0 \tag{III.17}$$

an estimate of Σ is

$$\hat{\Sigma} = \frac{1}{n-1} \sum_{i=1}^{n} (x_{mi} - \bar{x}_m)(x_{mi} - \bar{x}_m)' \tag{III.18}$$

from this we infer that

$$t^2 = \frac{n}{n+1}(x - \bar{x}_m)'\hat{\Sigma}^{-1}(x - \bar{x}_m) \sim T_{3,n-1}^2 \tag{III.19}$$

and therefore

$$\frac{n-3}{3(n-1)}t^2 \sim F_{3,n-3} \qquad \text{(III.20)}$$

If ς is the level of significance chosen then the probability that

$$\frac{n-3}{3(n-1)}\frac{n}{n+1}(x-\bar{x}_m)'\hat{\Sigma}^{-1}(x-\bar{x}_m) < F_{3,n-3,\varsigma} \qquad \text{(III.21)}$$

is $(1-\varsigma)$, where ς is the level of significance and is selected to be less than or equal to 0.05.

From the above equation we obtain

$$(x-\bar{x}_m)'\hat{\Sigma}^{-1}(x-\bar{x}_m) < \frac{3(n^2-1)}{n(n-3)}F_{3,n-3,\varsigma} \qquad \text{(III.22)}$$

or

$$x'\hat{\Sigma}^{-1}x < \bar{x}_m'\hat{\Sigma}^{-1}\bar{x}_m \frac{3(n^2-1)}{n(n-3)}F_{3,n-3,\varsigma} \qquad \text{(III.23)}$$

The right hand side of the above inequality is computed from the clean data after the sampling and formulation of the classifier. Every time we need to classify a new panel then we compute its vector x of r,g,b, modes and after computing the left hand side $x'\hat{\Sigma}^{-1}x$ of the above equation we compare it with the right hand side. If it satisfies the above inequality then we classify it as clean otherwise we trigger the cleaning mechanism to clean it.

4   Conclusion

Solar panels in the arid climates produce a great deal of energy because the sun shines either throughout the year, or close to 350 days of the year. Sandstorms, small animals, and birds, depositing their droppings are responsible for the reduction of the energy produced by the panels. Here we introduced an algorithm that takes as input the panel picture and decides if the panel is clean or needs cleaning. This project is relatively new. The algorithm we described in this paper is still being tested. So far our classification algorithm seems to be working. As the panels get dusty or subjected to bird droppings but they do not have enough deposits to pass our cut off point we see that the energy drop is within the normal fluctuation energy of clean panels. As they get dustier to the point that the inequality III.21 is not being satisfied then we see that there is enough drop in energy that by investing some energy to clean the panel will increase the energy output enough so that we can recover the energy invested by the panel within a few hours and have an energy gain until the next cleaning is required. The mean time between cleaning could be several months or less than a day depending on the weather conditions, the day of the month, the month of the year, and even the sun spot periodicity.

REFERENCES

[1].   C. E. Bohren and A. B. Fraser, "Colors of the Sky," The Physics Teacher, no. 5, 1985, pp. 269-272.

[2].   H. C. V. D. Hulst, "Light Scattering by Small Particles," John Wiley and Sons Inc., New York, 1957.

[3].   P. E. Haralabidis, and C. Pilinis, "Skylight Color Shifts due to Variations of Urban-Industrial Aerosol Properties: Observer Color Difference Sensitivity Compared to a Digital Camera," Aerosol Science and Technology, vol. 8, no 42, pp. 658-673, 2008.

[4].   E. Zamora Ramos, "Using Image Processing Techniques To Estimate The air Quality," Journal of Mcnair Scholars Institute, 6th Edition, pp. 189-194.

[5].   S.-C Tsay, G. L. Stephens and T. J. Greenwald,, "An Investigation of aerosol Microstructure on Visual air Quality," atmospheric environment, vol. 25A, no 5/6, pp. 1039-1053, 1991.

[6].   R. v. Hogg and A. T. Craig, "Introduction to mathematical Statistics,", Fourth Edition, Macmillan Publishing Co., pp. 1-438, 1978.

[7].   S. Theodoridis, K. Koutroumbas, "Pattern recognition," fourth edition, Academic Press, pp. 1-961.

A Novel Method of Brain Tumor Segmentation using Extensive Feature Set

M.Y.Bhanumurthy[1] and Koteswararao Anne[2]

[1]*Dept.of ECE, Vasireddy Venkatadri Institute of Technology, Guntur-522006, A.P, INDIA;*
[2]*Dean Academics, V.R.Siddhartha Engineering College, Vijayawada, A.P, INDIA;*
mybhanu@gmail.com; raoanne@gmail.com

ABSTRACT

Brain tissue Segmentation from the MRI images is having significance in the medical research field. The accurate Segmentation of the normal as well as the abnormal tissues is the complex assignment in this process. Because of the inconsistency and difficulty of abnormal tissues, MRI Brain Image Segmentation turned into more hard procedure. In this paper, a technique is proposed for segmenting the abnormalities such as Tumor and Atrophy in the MRI Brain images. (1) Feature extraction (2) Classification (3) Segmentation are the three stages offered in this work. At first, the features such as energy, entropy, homogeneity, contrast and correlation from MRI Brain Images are extracted. Next, by utilizing Neuro-Fuzzy classifier, the Classification process is carried out and for this process, the feature set is specified as the input. From the outcome of Classification, the images are categorized into normal as well as abnormal. The further procedure Segmentation is performed according to this outcome only. The abnormal MRI images are segmented into abnormal tissues like Tumor and Atrophy using Region Growing method. Utilizing MATLAB platform the implementation of the proposed technique is made. The experimentation is carried out on the MRI Brain Images by BrainWeb data sets. The performance of our proposed technique is assessed with the help of the metrics namely FPR, FNR, Specificity, Sensitivity and Accuracy. Therefore, using our proposed technique with enhanced classification, the abnormal tissues of MRI Brain images are segmented accurately.

Keywords:- Segmentation, Classification, Neuro-Fuzzy Logic, Normal and abnormal tissues, Region Growing method.

1 Introduction

The brain is the frontal most part of the central nervous system. It forms the Central Nervous System (CNS) along with the spinal cord. The Cranium, a bony box in the skull guards it. Because of our brain in practical we do lots of things like, to think, act, reason, walk, talk, the list is never-ending. Brain Tumors are one of the syndrome caused in the brain. In an (MR) images processing. As, in a number of neurological disorders like multiple sclerosis (MS) and Alzheimer's disease, the volume changes in total brain, WM, and GM can give major notification about neuronal and axonal loss [16].

The remaining of the paper is prepared as follows: After Introduction, the next section surveys several works which describes various techniques for segmenting the tissues in the MRI Brain images. Section 3 explains our proposed technique of segmenting the abnormal tissues of Brain MRI (BMRI) images. The outcomes regarding the performance of our proposed work are specified in the Section 4 and as a final point; our paper is summed up with the conclusion part in Section 5.

2   Literature Survey

Researchers proposed for many researches for the brain image segmentation. A short access of several researches is offered here. Arnaldo Mayer and Hayit Greenspan [17] have offered an automated segmentation framework for brain MRI volumes based on adaptive mean-shift grouping in the joint spatial and intensity feature space. The technique was authorized both on simulated and real brain datasets, and the outcomes were compared with state-of-the-art algorithms. The benefits over intensity based GMM EM schemes as well as additional state-of-the-art techniques were established. Moreover they proved that by means of the AMS framework, segmentation of the normal tissues is not degraded by the presence of abnormal tissues. The algorithm gave good outcomes on noisy and biased data while only a rudimental bias field improvement part executed and no spatial prior was extracted from an atlas. And thanks to the adaptive mean-shift ability to work with non-convex clusters in the joint spatial intensity feature space and also the mean-shift noise smoothing behavior.

Mert R. Sabuncu et al. [18] have examined a generative model that guides to label fusion style image segmentation techniques. They originated several algorithms that merge transmitted training labels into a single segmentation estimate in the proposed framework. An expert gave a dataset of 39 brain MRI scans and equivalent label maps and we analytically compared these segmentation algorithms with Free Surfer's broadly-used atlas-based segmentation tool. Their outcomes established that the proposed framework yields an accurate and robust segmentation tools that are employed on large multi-subject datasets. They utilized one of the enhanced segmentation algorithms to calculate hippocampal volumes in MRI scans of 282 subjects, in a second experiment. A assessment of these measurements across clinical and age groups signifies that the proposed algorithms were adequately sensitive to detect hippocampal volume variations related with earlier Alzheimer's disease and aging.

By utilizing a subject-specific tissue probabilistic atlas produced from longitudinal data, Feng Shi et al. [19] have offered a framework for presenting neonatal brain tissue segmentation. Proposed method has received the benefit of longitudinal imaging study in their system, i.e., by means of the segmentation outcomes of the images obtained at a late time to direct the segmentation of the images obtained at neonatal stage. Compared to the two population-based atlases the testing outcomes revealed that the subject-specific atlas has better performance. And moreover the proposed algorithm attained comparable performance as manual raters in neonate brain image segmentation. By attaining optimal segmentation results in a broad range of 0.3–0.6, the atlas sharpness parameter has been shown robust appearance. For the selection of late time-point image, the segmentation accuracy remains alike when the atlas was developed by either one-year-old or two-year-old image.

Juin-Der Lee et al. [20] have offered the most statistical segmentation methods in the literature and have presumed that either the intensity allocation of every tissue variety was Gaussian, or the logarithmic transformation of the raw intensity was Gaussian. As an alternative of setting up further classes to model "mixels," they proposed a power transformation approach to carry out automatic segmentation of brain MR images into CSF, GM, and WM. By instinct it was understandable that the familiar Box-Cox power transformation model was capable to give a statistically significant and helpful solution to proposed difficulty. To include both Gaussian intensity distributions as well as non-Gaussian distributions, the shape parameter utilized to widen the traditional Gaussian mixture models. And the parameters can be expected by means of the EM algorithm. They authorized the approach against four real and simulated datasets of normal brains from the IBSR and BrainWeb.

Testing's on real data from the IBSR have shown that compared with other techniques utilized presently, the proposed approach attains higher Jaccard indexes. The power transformation approach maintains the simplicity of the Gaussian mixtures, and in addition it has the prospective to simplify the multivariate versions personalized for segmentation by means of multi-modality images.

Dalila Cherifi et al. [21] have illustrated normal tissue's recognition than tumor extraction (applied for GBM and MS diseases). To detach the abnormal tissues they have offered brain recognition techniques. Based on thresholding utilized for tumor extraction (GBM and MS diseases) they have proposed and applied the technique. They have originated that the local thresholding provides a good outcomes comparing with the others. They have accomplished that when they merge median filter, local thresholding and post processing in such a way that the resultant algorithm is tougher. For tissue recognition and tumor extraction they have executed categorization based on EM segmentation technique. Comparing with thresholding particularly for detecting the small regions of necrotizing tissue which was inside Anaplastic cells (pseudo-Palisading necrosis) for GBM tissue, proposed technique provided us better outcomes; and it mainly for the reason that of parameters that utilized in this algorithm.

Nagesh Vadaparthi et al. [22] have offered a paper in which particular cases like Acoustic neuroma, it was presumed that there was an option of hearing loss, dizziness and other symptoms associated to brain. Surgery can cure various acoustic neuromas. Hence, it was required to segment the image more correctly, which assisted to recognize the damaged tissues to be repaired and can be corrected by surgery. And so a new novel segmentation algorithm based on Skew Gaussian distribution was proposed in proposed paper, which assisted to recognize the tissues more correctly. Because of the basic structure of Skew Gaussian distribution it was suitable for symmetric and asymmetric distribution. The performance evaluation was succeeded by utilizing quality metrics. The outcomes proved that, proposed developed algorithm outperforms the existing algorithm. Various models were exploited to recognize the diseases, although due to the utility of non-ionizing radiation, MRI brain segmentation has achieved popularity over the other models.

Usually noise is generated by equipments, environment and also the performance of operator in MRI Brain images which creates serious incorrectness in the outcome of Segmentation procedure. Several of the unverified techniques did not deal with the intensity and in-homogeneity artifacts. And also the managed techniques undergo with the shortcomings of manual intervention for providing a priori notification. Pathological tissues demonstrate inconsistency in their structures. The shape of these tissues is deformable, the location of them across the patients may differ extensively, and also their characteristics of texture and intensity might vary. These difficulties of the existing schemes are generally un-solvable. A few techniques do not consider the large deformation of brain structures. The practice of brain atlas might show the way to false learning, though such deformations occurs.

3   Proposed Methodology

Initially, the BMRI images are given as input to our proposed work and the feature sets are extracted from these input images. From these feature sets, the images are classified into two kinds of tissues – normal and abnormal using the Neuro-Fuzzy classifier. Then the abnormal tissues Tumor and Atrophy are segmented using Region Growing Method. The proposed work is illustrated in Fig. 1.

3.1 Phases of Neuro-Fuzzy based segmentation

For our proposed method, to segment the BMRI images effectively, the three phases are presented which are as follows:

I. Feature set Extraction

II. Neuro-Fuzzy classifier based Classification

III. Classified tissue's Segmentation

3.1.1 Phase I: Extraction of extensive feature sets

In order to classify the given Brain MRI images, the features from these MRI images are initially extracted. In our work, the statistical features such as Energy, Entropy, Homogeneity, Contrast and Correlation are extracted from these input BMRI images.

Energy

Energy is also called as uniformity. Within [0,1] the range of energy is presented. The value of energy for a constant image is 1. The equation for finding energy is,

$$E_g = \sum_{i,j} p(i,j)^2 \tag{1}$$

where, $p(i,j)$ is the pixel value at the point i,j of the BMRI image of size $M \times N$.
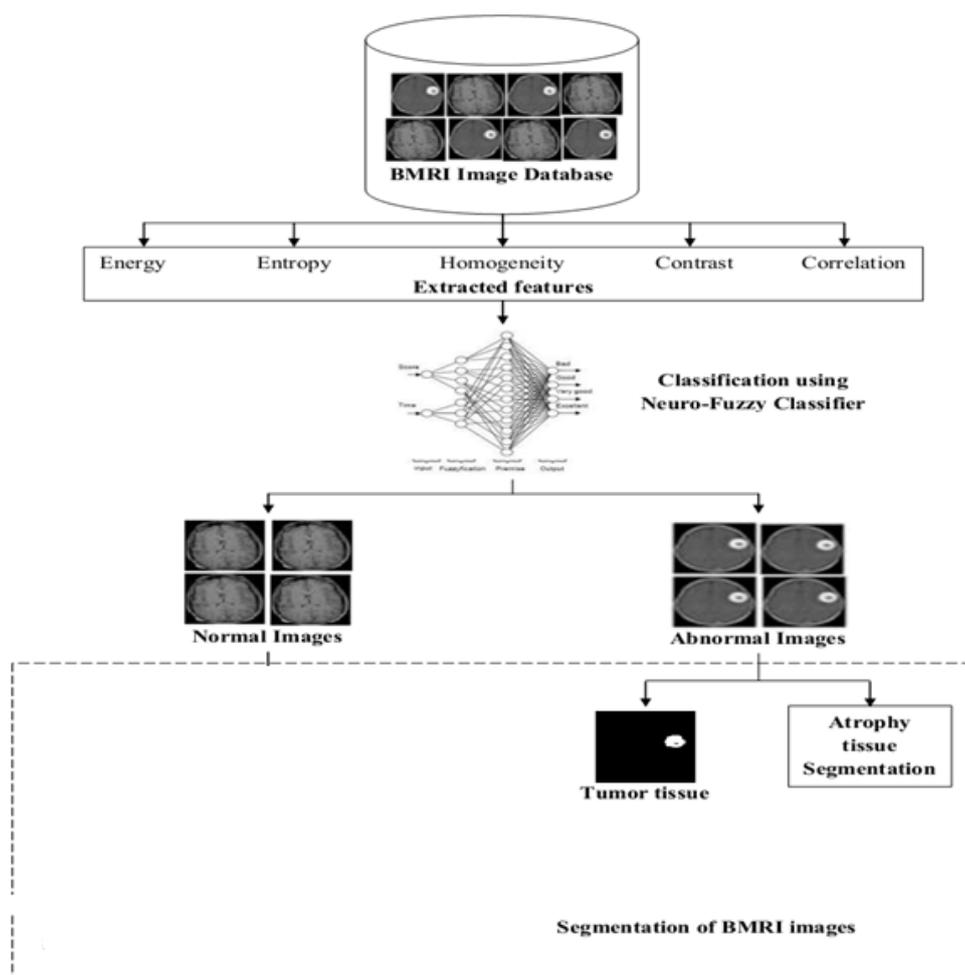

Figure 1: Proposed Neuro-Fuzzy based segmentation block diagram

Entropy

Entropy helps to characterize the texture of the BMRI image and to find out the distribution variation in a region of the image. Entropy is calculated as follows,

$$E_p = \sum_{k=0}^{G-1} Pb_k \left(\log_2 Pb_k\right) \tag{2}$$

where, Pb_k is the probability of kth gray level and the kth gray level is calculated using $\dfrac{Z_k}{M \times N}$. In this, Z_k represents the total number of pixels in the image with kth gray levels. G indicates the total number of gray levels.

Homogeneity

Homogeneity provides the closeness of the elements. It has the range of [0,1]. It is computed as,

$$H_m = \sum_{i,j} \frac{p(i,j)}{1+|i-j|} \tag{3}$$

Contrast

The intensity contrast between a pixel of an image and the neighbor of that pixel throughout the whole image is defined by this Contrast measure. For a constant image, the contrast is set as 0. It is specified as,

$$C_n = \sum_{i,j} |i-j|^2 \, p(i,j) \tag{4}$$

Correlation

It tells about the correlation between a pixel and its neighbor over the whole BMRI image. Its range is [-1,1]. If an image has the value of correlation as 1 means, then it indicates the perfectly positively correlated image and if it is -1 means, then it shows the image is perfectly negatively correlated. The correlation of a constant image is not a number.

$$C_r = \sum_{i,j} \frac{(i-\mu_i)\,(j-\mu_j)\,p(i,j)}{\sigma_i \sigma_j} \tag{5}$$

where, μ_i, μ_j, σ_i, σ_j are the means and standard deviations of the partial probability density functions P_i, P_j.

Mean, $$\mu = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} p(i,j)^2$$

Standard Deviation, $$\sigma = \sqrt{\frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} (p(i,j)-\mu)^2}$$

Variance, $$Var = \sqrt{\sigma}$$

Thus the feature correlation of the images is calculated using the mean and variance equations. Hence, all the feature sets $E_g, E_p, H_m, C_n, \text{ and } C_r$ are extracted from the input BMRI images directly.

3.1.2 Phase II: Classification using Neuro-fuzzy classifier

The BMRI images are classified using the Neuro-Fuzzy classifier. The extracted features $E_g, E_p, H_m, C_n,$ *and* C_r are given as the input to the Neuro-Fuzzy Classifier for classifying all the given BMRI images into 2 classes such as Normal BMRI images and Abnormal BMRI images. The Neuro-fuzzy system has a three-layered architectural design; the following diagram fig. 2 shows the basic structure of the neuro-fuzzy classifier system. Neuro-Fuzzy classifier is a fuzzy based system that is trained by a learning algorithm derived from Neural Networks. The learning algorithm only performs on the local information and provides the local modifications in the fuzzy system. In general, a neuro-fuzzy system generates very powerful solutions instead of using the system components individually.
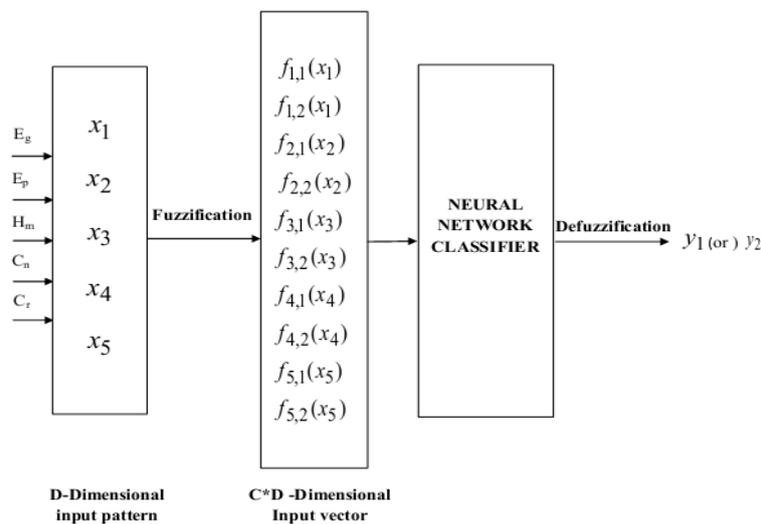

Figure 2: Architecture of Neuro-Fuzzy classifier

Fuzzification

The input values are the extracted features $E_g, E_p, H_m, C_n,$ *and* C_r, which are received by the system as the input and then these input feature values are fuzzified using membership functions (MF) that facilitates the membership of each features to different classes. The hidden and inter-related information are extracted from the features to the classes through the MF, which leads to get more accuracy of the classification phase using Neuro-fuzzy classifier. The membership matrix comprises with 5 rows and 2 columns, in which the number of rows is equal to the number of features and the number of columns is equal to the number of classes.

The membership matrix $f_{d,c}(x_d)$ produced, describes the degree of belonging of different features (D) to different classes (C).

Where, x_d - d^{th} feature value of pattern X.

d - 1, 2, …, D, here number of features is 5.

c - 1, 2, …, C, here number of classes is 2.

The representation of pattern is as follows,

$$X = [x_1, x_2, x_3, x_4, x_5]^T \qquad (6)$$

In this a π–type MF is used as the membership function to classify the images. It is a bounded function having a shape similar to that of the Gaussian/exponential function. The π–type MF has fuzzifier (m) as the parameter that can be tuned corresponding to the need of the problem. This controls the generalization capability by choosing a proper value of the fuzzifier m and gives more flexibility for classifying the images. The steepness of the Gaussian function is controlled by varying the fuzzifier value, which is defined as follows,

$$
\pi(X;a,r,b) = \begin{cases}
0, & if\ X \leq a \\[2mm]
2^{m-1}\left[\dfrac{(X-a)}{(r-a)}\right]^m, & if\ a < X \leq p \\[4mm]
1 - 2^{m-1}\left[\dfrac{(r-X)}{(r-a)}\right]^m, & if\ p < X \leq r \\[4mm]
2^{m-1}\left[\dfrac{(X-r)}{(b-r)}\right]^m, & if\ r < X \leq q \\[4mm]
1 - 2^{m-1}\left[\dfrac{(b-X)}{(b-r)}\right]^m, & if\ q < X < b \\[4mm]
0, & if\ X \geq b
\end{cases}
\tag{7}
$$

The value r is the center of MF, and $r = \dfrac{(p+q)}{2}$, in which p and q are the two crossover points.

The membership function after the fuzzification process is expressed for a pattern X as follows,

$$
F(X) = \begin{bmatrix}
f_{1,1}(x_1) & f_{1,2}(x_1) \\
f_{2,1}(x_2) & f_{2,2}(x_2) \\
f_{3,1}(x_3) & f_{3,2}(x_3) \\
f_{4,1}(x_4) & f_{4,2}(x_4) \\
f_{5,1}(x_5) & f_{5,2}(x_5)
\end{bmatrix}
\tag{8}
$$

All rows and columns in the membership matrix are cascaded and converted into a vector by this cascading. This generated vector V_i is given as the input to the Neural Network (NN).

Neural Network

In this, Feed Forward Multi-layer Perceptron classifier is used which has three layers such as input layer, hidden layer and output layer.

The total number of input nodes of the NN is equal to the product of the number of features and classes. In this paper the product of 5 features and 2 classes is 10, which is the number of input nodes of the NN. The total number of output nodes from the NN is same as that of the number of classes, and here 2 output nodes are generated from the NN. The total number of hidden nodes is equal to the square root of the product, of the number of input nodes and output nodes. The structure for the Neural Network is given in figure 3.
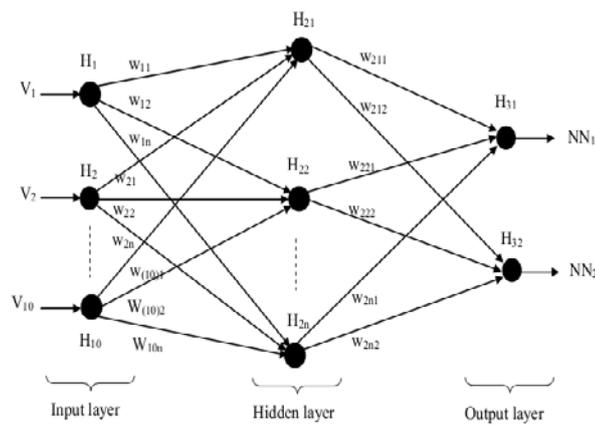
Figure 3: Neural Network classifier

Defuzzification

Then the defuzzification process is carried out on the output nodes of NN, by performing a MAX (maximum) operation. The output is a single value, y_1 or y_2 for a given BMRI image. From this value, we can able to classify whether the given input BMRI image is normal or abnormal.

3.1.3   Segmentation of classified tissues

Utilizing Neuro-fuzzy classifier the BMRI images are classified and after that the images are comprised only in any of the two different images normal and abnormal or pathological images. From the abnormal images, tumor and atrophy are segmented.

Segmentation of Tumor

For segmenting the tumor tissues, the classified abnormal images are utilized. Region Growing Method (RGM) is used for the segmentation of tumor tissues. RGM is one of the image segmentation techniques, in which the first seed points are chosen from the abnormal images. For the neighboring pixels in the first seed points, the neighbor pixels are checked out to verify the neighboring pixels are located within the region or not [23]. Accordingly, the tumor tissue part is segmented from the abnormal image and the resultant image is I_T .

Segmentation of Atrophy

By using the White Matter and Gray Matter ratio as the whole brain, the degree of atrophy is established from the abnormal image. This is the way to identify the atrophy in the beginning stage.

Atrophy Ratio (AR)

The Atrophy Ratio (AR) is computed by comparing the White Matter and Gray Matter with the whole size of the brain, which includes White Matter, Gray Matter and Cerebro-Spinal Fluid. For a BMRI image, the atrophy ratio is calculated as follows:

$$AR = \frac{(WM + GM)}{(WM + GM + CSF)} \tag{9}$$

In this above equation (9), WM, GM and CSF specify the region area of White Matter, Gray Matter and Cerebro-Spinal Fluid respectively. We can tell the atrophy level in the brain, according to the ratio of Atrophy. It is studied that the existence of atrophy is in high level, if the ratio value is small. From the Atrophy ratio, the diseases like multiple sclerosis, Alzheimer's disease, Pick's disease, senile dementia, vascular dementia, stroke, etc. [24] are identified.

Atrophy Factor (AF)

Atrophy Factor is a measurement that helps out to study the Atrophy Rate among two consecutive check-ups from the BMRI images for the same patient within a specific time interval. The AF is computed as specified in the below equation (10).

$$AF = \frac{A_{t1} - A_{t2}}{A_{t1}} \qquad (10)$$

In the above equation (10), A_{t1} and A_{t2} are the two consecutive MRI checkups within a specific time interval for the Atrophy of whole brain. In general, the difference in time is one year. We can say that no Atrophy is obtainable throughout this year, if the resultant AF value is too small or almost zero. If the value of AF is improved, then the level of Atrophy in brain is as well high. Therefore, the tumor and atrophy segmentation are efficiently made from the categorized abnormal BMRI images.

4 Results and Discussions

Our proposed Neuro-Fuzzy based segmentation for the effective segmentation of tumor and atrophy is implemented using the MATLAB platform on the Brain MRI images from the dataset. The data set description is given below in detail.

4.1 Dataset Description:

BrainWeb dataset is utilized with different BMRI images for our proposed work. Based on standard tissue segmentation mask, BrainWeb datasets give MRI brain images with unreliable image quality. The datasets are too based on an anatomical structure of a normal brain, which results from the tasks of registering and preprocessing of 27 scans from the same individual with segmentation. Different kinds of tissues are well identified in this dataset, both the types of tissue memberships "fuzzy" and "crisp" are assigned to each voxel. The sample Brain MRI images from the BrainWeb data set are specified in the figure 4 given below.
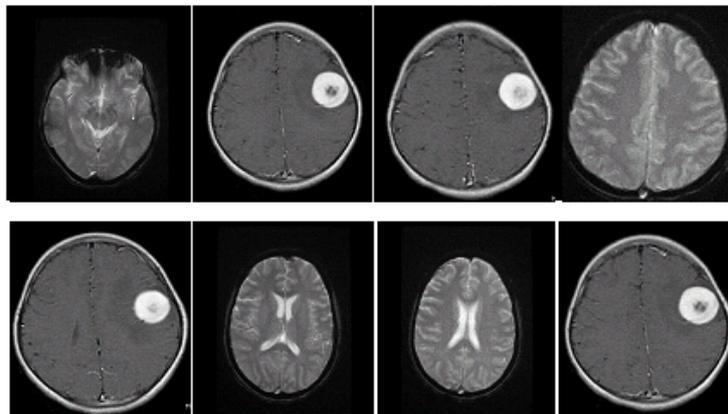

Figure 4: Sample BMRI images from dataset

Our proposed work is estimated by means of 23 BrainWeb MRI images. 9 images are normal and the remaining 14 images are abnormal among 23 MRI images. At first, the BMRI images from this dataset is taken and offered to the procedure of our proposed NFBIS. Five of the statistical characteristics from these BMRI images are extracted. Then for the categorization of images these extracted features are utilized. In order to categorize the particular images into normal and abnormal images, Neuro-Fuzzy classifier is utilized as the classifier in this proposed NFBIS work. The classified normal and abnormal images are specified in the   figure 5 given below.
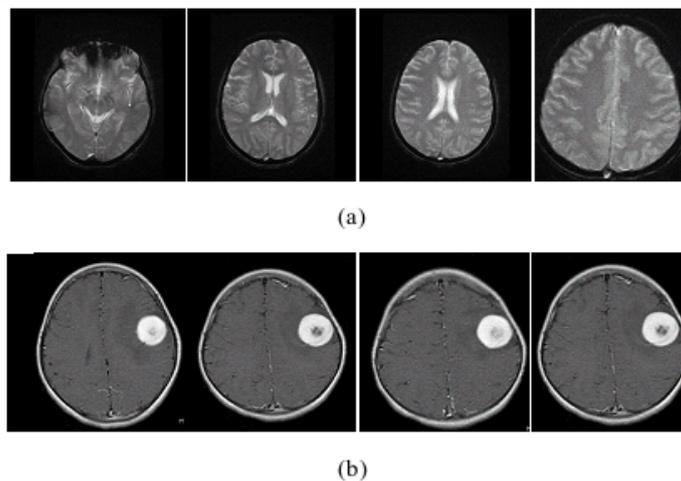
(a)

(b)

Figure 5: Normal and Abnormal Images

From the abnormal BMRI images, the abnormal tissues Tumor and Atrophy are segmented. By means of Region Growing Method, the segmentation of tumor is performed. The images of tumor tissues after segmentation are specified in the following figure 6.
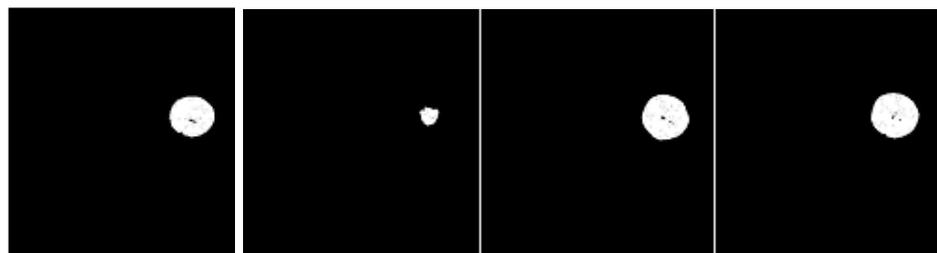

Figure 6: Segmentation of abnormal images

Table I provides the tabular values for the accuracy of abnormal tissue atrophy segmentation.

Table I: Atrophy level results for various images

Images	Atrophy level (in %)
Image-1	0.790629155
Image-2	0.652814507
Image-3	0.875238649
Image-4	0.865906419
Image-5	0.966725302

The image-5 has the extremely high atrophy level (96.67%), which shows that the existence of atrophy is in low level. Next the existence of atrophy level is low level than the image 5 for the images 3 and 4, as the values are 87.52% and 86.59%, respectively. The existence of atrophy is high in image 2 and 1, since both have the values 65.28% and 79.06% respectively.

4.2. Performance Evaluation:

By utilizing the performance measures namely False Positive Rate, False Negative Rate, Sensitivity, Specificity and Accuracy, the performance of the system is estimated. The basic count values such as True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) are used by these measures. Both the categorization of normal and abnormal images and the segmentation efficiency of abnormal images are examined by our proposed work, which are clarified in detail in the next sections.

4.2.1 Results of Classification Evaluation:

The BrainWeb images include both the normal and abnormal images in our work. These images are categorized into normal and abnormal individually by the procedure of Neuro-Fuzzy classifier. The efficiency of the classifier is examined by the metrics False Positive Rate, False Negative Rate, Sensitivity, Specificity and Accuracy. The explanation of TP, TN, FP, FN values for the categorization of normal and abnormal images is specified in the table II given below.

Table II: Description of TP, TN, FP, FN values for the classification of normal and abnormal images

Description	Classified as normal image	Classified as abnormal image
Actually normal image	TP	FN
Actually abnormal image	FP	TN

False Positive Rate (FPR)

The percentage of cases where an image was classified to normal images, but in fact it did not.

$$FPR = \frac{FP}{FP+TN} \tag{11}$$

False Negative Rate (FNR)

The percentage of cases where an image was classified to abnormal images, but in fact it did.

$$FNR = \frac{FN}{FN+TP} \tag{12}$$

Sensitivity

The proportion of actual positives which are correctly identified is the measure of the sensitivity. It relates to the ability of test to identify positive results.

$$Sensitivity = \frac{Number\ of\ true\ positives}{Number\ of\ true\ positives + Number\ of\ false\ negatives} \times 100 \tag{13}$$

Specificity

The proportion of negatives which are correctly identified is the measure of the specificity. It relates to the ability of test to identify negative results.

$$Specificity = \frac{Number\ of\ true\ negatives}{Number\ of\ true\ negatives + Number\ of\ false\ positives} \times 100 \tag{14}$$

Accuracy

We can compute the measure of accuracy from the measures of sensitivity and specificity as specified below.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \tag{15}$$

The subsequent table III explains the categorization efficiency outcomes for the normal and abnormal images with various metric values.

Table III: Effectiveness of classification results using Neuro-Fuzzy classifier for the normal and abnormal images

Metrics	Values
TP	8
TN	14
FP	0
FN	1
FPR	0
FNR	6.667
Sensitivity	88.9%
Specificity	100%
Accuracy	95.65%

In our proposed work, we can establish the efficiency of categorization for the normal and abnormal images by means of Neuro-Fuzzy classifier from the above table II. False Positive Rate and False Negative Rate values are 0 and 6.667, respectively, which explains that our proposed work has low error rate in categorizing images. Properly categorized percentages of normal images are specified by Sensitivity. Neuro-fuzzy classifier offers very high (88.9%) values for the metric sensitivity, in which only one of the normal image is categorized as abnormal. In addition, Specificity is another metric that specifies the percentage of abnormal images properly categorized. The classifier provides 100% specificity rate by categorizing the entire actual abnormal images into abnormal images in our work. The highest value in sensitivity and specificity and also the lowest value in the error rates False Positive Rate and False Negative Rate open a mode to raise the categorization correctness outcome with the value 95.65%. Therefore we can show that in proposed work categorization of BMRI BrainWeb images offers high classification accuracy.

4.2.2   Results of Segmentation Evaluation:

The abnormal tissues namely tumor and atrophy are segmented from the BMRI images of BrainWeb. The tumor segmentation outcomes are tabularized in the following table IV.

Table IV: Tumor segmentation results for various images

Images	TP	TN	FP	FN	FPR	FNR	Sensitivity (in %)	Specificity (in %)	Accuracy (in %)
Img 1	5405	255045	1694	864	0.0065981	0.137821	86.217	99.34018	99.0274
Img 2	5405	255045	1694	864	0.0065981	0.137821	86.217	99.34018	99.0274
Img 3	3870	252820	5454	102601	0.0211171	0.963652	3.634	97.88828	70.3751
Img 4	5405	255045	1694	864	0.0065981	0.137821	86.217	99.34018	99.0274
Img 5	3870	252820	5454	102601	0.0211171	0.963652	3.634	97.88828	70.3751

Corresponding graph of table IV is designed in Figure 7 with different BMRI images for the Tumor segmentation. The evaluation outcomes explains whether our proposed work is worked efficiently or not for the tumor segmentation.
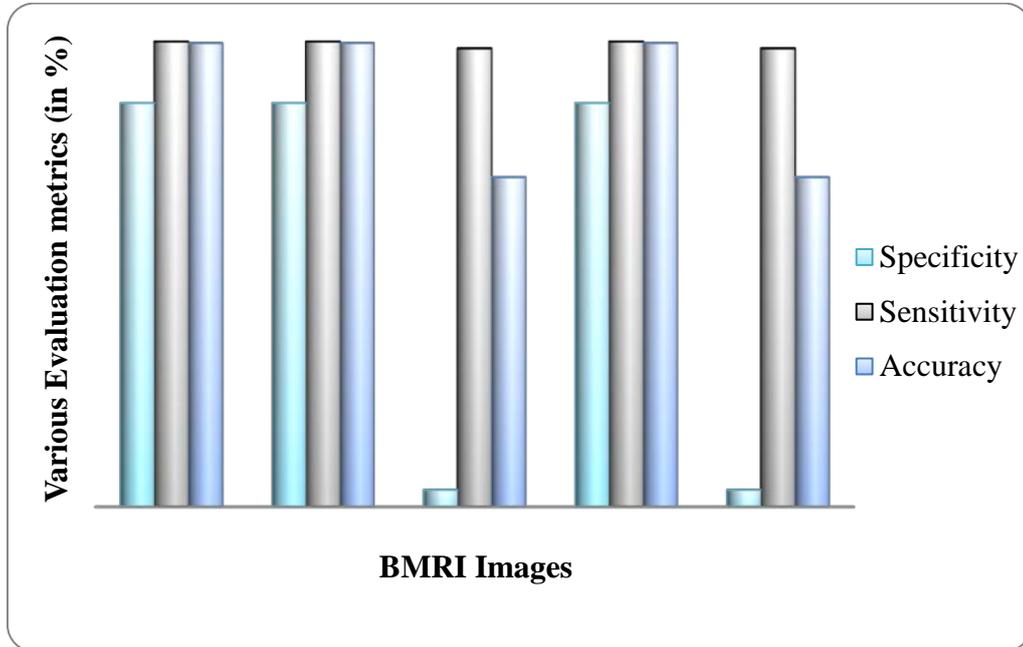
Figure 7: Sensitivity, Specificity and Accuracy results of Tumor segmentation

The outcomes of tumor segmentation are viewed with the metric values from the table IV and Figure 7. The abnormal tumor tissue segmentation of our proposed work presents improved accuracy outcomes of 99.027%, 99.027%, 70.375%, 99.027% and 70.375% for the images 1, 2, 3, 4 and 5, respectively. The outcomes of specificity give high values of 99.34%, 99.34%, 97.88%, 99.34% and 97.88% for the images 1, 2, 3, 4 and 5. The raise in the specificity values builds the tumor segmentation very accurately. The sensitivity values for the images 1, 2 and 3 are high (86.217%) and for the images 3 and 5 are very low (3.634%). The low value in sensitivity is not a difficulty to obtain accuracy with improved outcome. In addition, our work provides only very low values for the error rates FPR and FNR. It's a great benefit for our work to get improved accuracy of segmentation. Correct tumor parts were only segmented from the abnormal BMRI images of our proposed work. This can be recognized from the high accuracy outcomes of tumor segmentation.

4.2.3   Comparative Analysis for our proposed work with the existing works:

For the categorization of normal and abnormal brain images our proposed work makes use of Neuro-Fuzzy classifier. We can establish that our proposed work helps to attain very good accuracy for the categorization of images utilizing Neuro-Fuzzy classifier from the above sections. And also we can establish this categorization accuracy outcome by comparing other classifiers. We have utilized Artificial Neural Network and Fuzzy C-Means for our comparison in our work. The comparison outcomes are presented in the following table V.

Table V: Comparison results for the image classification with other classifiers

Metrics	Fuzzy C-Means	Artificial Neural Network	Neuro-Fuzzy in our proposed work
TP	0	1	8
TN	14	9	14
FP	0	5	0
FN	9	8	1
FPR	0	0.3571	0
FNR	0.3913	0.4706	6.667
Sensitivity (in %)	0	11.11	88.9
Specificity (in %)	100	64.29	100
Accuracy (in %)	60.87	43.48	95.65

Below specified figure 8 explains the comparison outcomes of the classifiers for the BMRI image categorization with different metrics. The improved accuracy outcomes of categorization of BMRI images into normal and abnormal images are presented by our proposed work. In comparison with the classifier Neuro-Fuzzy, both the Fuzzy C-Means and Artificial Neural Networks gives very less accuracy values for the categorization of images. The sensitivity for the Fuzzy C-Means and Artificial Neural Networks are 0% and 11.11%, which is low in compared with our classifier, Neuro-Fuzzy 88.9%. The specificity is 100% for our classifier and for the Fuzzy C-means classifier. However the accuracy is 95.65% for our Neuro-fuzzy classifier and the fuzzy C-means and ANN contain only low categorization accuracy results of 60.87% and 43.48%, respectively. From these outcomes, it is known that by means of Neuro-Fuzzy classifier in our work provides very good for the categorization purpose as it gives improved accuracy outcomes. Therefore, our work shows that it is worth for the categorization and segmentation of BMRI images.
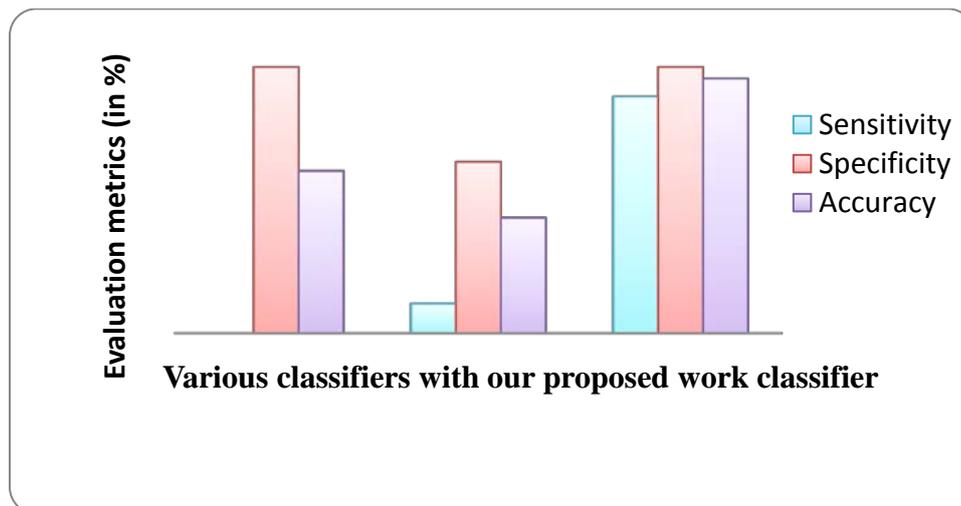

Figure 8: Comparison graph for the image classification with other classifiers

5 Conclusion

A Neuro Fuzzy based BMRI image segmentation technique with three phases – Feature Extraction, Classification and Segmentation was proposed in this paper. The features from the BMRI images were extracted and then specified to the Neuro-Fuzzy classifier. The classification of normal and abnormal images was made by this Neuro-Fuzzy classifier and these classified abnormal images were segmented efficiently by our proposed method. The testing was performed with the BrainWeb

images dataset. The performance measures False Positive Rate, False Negative Rate, sensitivity, specificity and accuracy were evaluated for our proposed method. The testing was prepared for establishing the accuracy of both the classification of images into normal and abnormal and segmentation of abnormal tissue like Tumor. The efficiency of the classification of images is very high by presenting very good accuracy outcomes and also the segmentation of Tumor offers very accurate outcomes. From the outcomes, we have showed that the Neuro-Fuzzy classifier utilized in our proposed work outperforms the other classifiers Fuzzy C-Means and ANN by facilitated very good accuracy of 95.65% in categorizing the images into normal and abnormal.

REFERENCES

[1]. P. Hagmann, J.P. Thiran, L. Jonasson, "DTI mapping of human brain connectivity: statistical fibre tracking and virtual dissection", NeuroImage, 2003.

[2]. M.C. Davidson, K. M. Thomas. and B. J. Casey, "Imaging the developing brain with fMRI", Mental Retardation and developmental disabilities research reviews, 2003.

[3]. V. A. Grau, U. J. Mewes, M. Alcaniz, "Improved watershed transform for medical image segmentation using prior information", IEEE Trans. on Medical Imaging, 2004, 23(4): 447-458

[4]. H Lv., K. H. Yuan, S. L. Bao, "An eSnake model for medical imaging segmentation", Progress in Natural Science, 2005.

[5]. D. L. Pham and J. L. Prince, "Adaptive fuzzy segmentation of magnetic resonance images," IEEE Trans. Med. Imag., 1999.

[6]. A. F. Goldszal, C. Davatzikos, D. L. Pham, M. X. H. Yan, et al, "An image processing system for qualitative and quantitative volumetric analysis of brain images," J. Comput. Assist. Tomogra., 1998.

[7]. Arnold J.B., Liow, J.-S., Schaper, K.A., et al., "Qualitative and quantitative evaluation of six algorithms for correcting intensity nonuniformity effects". NeuroImage, 2001.

[8]. R. Moller., R. Zeipelt. "Automatic segmentation of 3D-MRI data using a genetic algorithm, Medical Imaging and Augmented Reality", 2001. Proceedings. International Workshop on, 10-12 June 2001:278 – 281.

[9]. W. M.Wells, III,W. E. L. Grimson, R. Kikinis. "Adaptive segmentation of MRI data", IEEE Trans. Medical Imaging , 1996.

[10]. J. C. Bezdek, L.O. Hall, L. P. Clarke, "Review of MR image segmentation techniques using pattern recognition," Med. Phys., vol. 20, No. 4, pp. 1033-1048, 1993.

[11]. Suetens, E. Bellon, D. Vandermeulen, M. Smet, G. Marchal, J. Nuyts, L. Mortelman, "Image segmentation: methods and applications in diagnostic radiology and nuclear medicine," European Journal of Radiology, vol. 17, pp. 14-21, 1993.

[12]. A. Goshtasby, D. A. Turner, "Segmentation of Cardiac Cine MR Images for extraction of right and left ventricular chambers," IEEE sTrans. Med. Imag., vol. 14, No. 1, pp. 56-64, 1995.

[13]. D. Brzakovic, X. M. Luo, P. Brzakovic, "An approach to automated detection of tumors in mammograms," IEEE Trans. Med. Imag., vol. 9, No. 3, pp. 233-241, 1990.

[14]. J. F. Brenner, J. M. Lester, W.D. Selles, "Scene segmentation in automated histopathology: techniques evolved from cytology automation," Pattern Recognition, vol. 13, pp. 65-77, 1981.

[15]. K. Lim, A. Pfefferbaum, "Segmentation of MR brain images into cerebrospinal fl¬uid spaces, white and gray matter," J. Comput. Assist. Tomogr., vol. 13, pp. 588-593, 1989.

[16]. Zhang Y, Brady M, Smith S. "Segmentation of brain MR images through a hidden Markov random field model and expectation-maximization algorithm,". IEEE Trans Med. Imag., pp. 45–57, 2001.

[17]. Arnaldo Mayer and Hayit Greenspan, "An Adaptive Mean-Shift Framework for MRI Brain Segmentation", IEEE Transactions On Medical Imaging, Vol. 28, No. 8, August 2009.

[18]. Mert R. Sabuncu, B.T. Thomas Yeo, Koen Van Leemput, Bruce Fischl and Polina Golland, "A Generative Model for Image Segmentation Based on Label Fusion", Ieee Transactions On Medical Imaging, 2009.

[19]. Feng Shi, Yong Fan, Songyuan Tang, John H. Gilmore, Weili Lin, Dinggang Shen, "Neonatal brain image segmentation in longitudinal MRI studies", Elsevier Inc., 2009.

[20]. Juin-Der Lee, Hong-Ren Su, Philip E. Cheng*, Michelle Liou, John A. D. Aston, Arthur C. Tsai, and Cheng-Yu Chen, "MR Image Segmentation Using a Power Transformation Approach", IEEE Transactions On Medical Imaging, Vol. 28, No. 6, June 2009.

[21]. Dalila Cherifi, M.Zinelabidine Doghmane, Amine Nait-Ali , Zakia Aici,Salim Bouzelha, "Abnormal tissus extraction in MRI Brain medical images", IEEE, 2011.

[22]. Nagesh Vadaparthi, Srinivas Yarramalle, Suresh Varma Penumatsa, "Unsupervised Medical Image Segmentation On Brain MRI Images Using Skew Gaussian Distribution", IEEE-International Conference on Recent Trends in Information Technology, ICRTIT 2011.

[23]. Frank Y. Shih and Shouxian Cheng, "Automatic seeded region growing for color image segmentation", Journal of Image and Vision Computing, Vol. 23, pp. 877–886, 2005.

[24]. Rowayda A. Sadek, "An Improved MRI Segmentation for Atrophy Assessment", International Journal of Computer Science Issues, Vol. 9, No 2, pp. 569-574, May 2012.

Implementation of Image Processing Lab Using Xilinx System Generator

K. Anil Kumar and M. Vijay Kumar

Department of Electronics and Communication Engineering,
Gudlavalleru Engineering College, INDIA;
anilkumar.kanna3@gmail.com; vij4b7@gmail.com

ABSTRACT

The paper presents information on various image processing operations using Field Programmable Gate Arrays (FPGA). Processing of images on FPGA is complicated, since it needs separate architectures to process the image. To facilitate such operations, Matlab, Simulink and Xilinx system generator tools, which convert the image into suitable formats that are supported by FPGA, are used. XSG plays an instrumental role in generating VHDL/VERILOG code in tune with algorithms designed in Simulink. The generated code will be dumped into FPGA and then it performs operations on image. Use of XSG in image processing effectively reduces total design time of a system.

Keywords: Image, Simulink, FPGA.

1   Introduction

Over the past decades, the field of image processing has undergone a rapid evolution. Image processing has varied applications, computer vision, digital photography [8], Traffic load computation [7] etc. Current trends in digital camera technology have led to an increase in larger number of pixels being adjusted into smaller spaces. This result in an overall descends in the visual quality of images.

This paper particularly determinates on developing appropriate method to perform hardware implementation of various image processing algorithms that can be used in some applications. Image quality can be enhanced by creating image processing algorithms using Xilinx system generator such as, contrast stretching[8], edge detection[1],etc. these algorithms were mainly discussed in this paper. This paper aims at, (1) Implementing algorithms in MATLAB using Xilinx system generator (XSG) for specific role, (2) Generation of HDL code using Xilinx system generator token, (3) Hardware implementation of given algorithms on FPGA.

2   XSG Design Flow

For accomplishing Image processing task using FPGA, MATLAB [12], Simulink [6] and Xilinx system generator[13] tools are used. Simulink is a model based design environment integrated with Matlab. One of the block library provided by Simulink is Xilinx system generator. The System Generator token along with Xilinx has to be mapped to MATLAB. This agglomerate Xilinx Blockset to the Matlab, Simulink environment which can be directly used for building algorithms. The algorithms are developed for image negative, image enhancement, contrast stretching etc. using Xilinx Block set. These algorithms are simulated in Matlab, Simulink environment with appropriate simulation time. After obtaining the results, System Generator is configured for appropriate FPGA board. FPGA board used here is Virtex5 xc5vlx110t-3ff1136.

After compilation, programming file in VHDL has been created and can be accessed using Xilinx ISE. The module is checked for syntax check, synthesized and implemented on FPGA. The Xilinx System Generator can generate User constraints file (UCF) for testing architecture. Bit stream compilation is necessary to create an FPGA bit file which is inevitable for FPGA input. The Figure 1 depicts the system generator design flow.
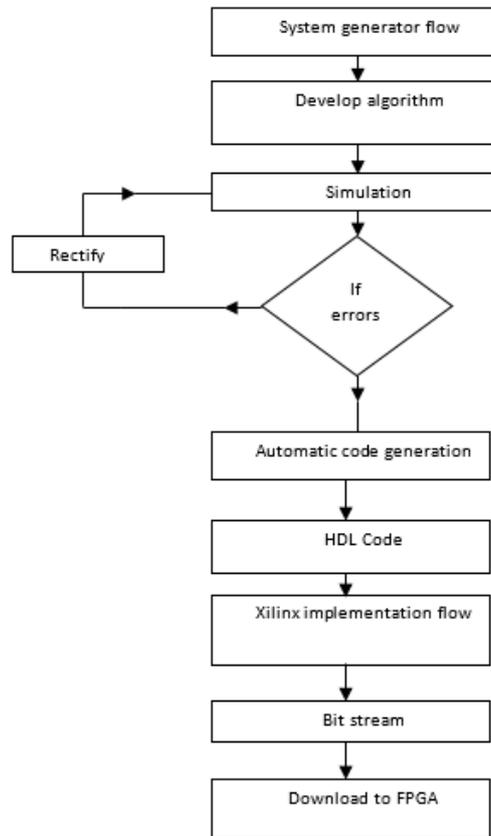

Figure 1: System Generator Design Flow
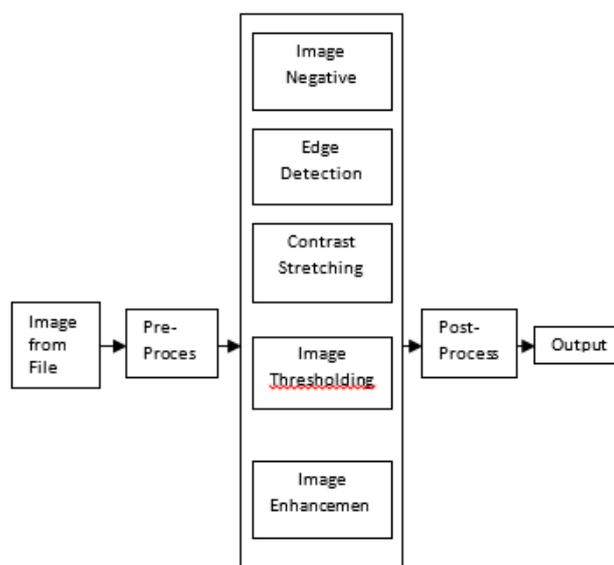
3   Design of Image Processing Lab on FPGA

Figure 2: Design of Image Processing Lab on FPGA

Before processing the image on Xilinx system generator, it should be applied to preprocessing [6] block, which converts 2-D (frame) into 1-D (serial), as Xilinx system generator doesn't able to process the image in 2-D.Various operations such as edge detection, negative generation etc. can be performed on the arrived 1-D (serial) data stream using Xilinx system generator. To obtain the output image properly post processing [6] must be done, which converts 1-D into 2-D.

3.1 Image Pre-Processing

Transpose   Convert 2-D to 1-D   Frame conversion
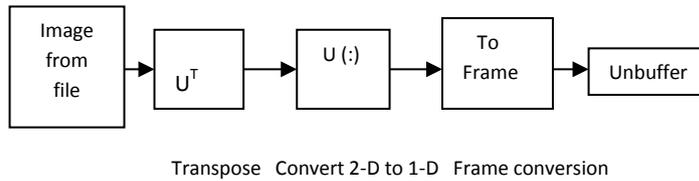
Figure 3: Pre-Processing Block Diagram

Image from file block reads the image from file. If the image is M-by-N array, the block outputs a binary or intensity image, where M and N are rows and columns. If the image is M-by-N-by-P array, the block outputs a color image, where M and N are rows and columns in each color plane, P. Then the transpose block transposes M-by-N matrix to N-by-M matrix. Convert 2-D to 1-D block converts the input data (2-D) to 1-D (serial) format. Later, Frame conversion block sets the output sampling mode to either frame based or sample based. Unbuffer block unbuffers an M-by-N input into a 1-by-N output. That is, inputs are unbuffered row wise.

3.2 Image Post-processing

Data type conversion   Convert 1-D to 2-D   Transpose
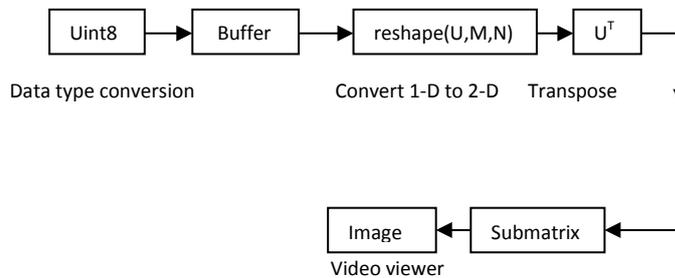
Figure 4: Post-Processing Block Diagram

Data type conversion block translates an input signal to the data type required for the Output data type parameter. The input can be any real or complex-valued signal. If the input is real, the output is real. If the input is complex, the output is complex. Buffer block performs frame-based processing. The block produces an output with a different frame size by redistributing the data in each column of the input. Buffering a signal to a larger frame size yields an output with a slower frame rate than the input. Convert 1-D to 2-D block converts the input data (1-D) to 2-D (frame) format. Sub matrix block extracts a contiguous sub matrix from the M-by-N input matrix.

4   Image Processing Algorithms

To implement image processing algorithms Xilinx system generator is used because of its ability to generate HDL code. Various algorithms like negative generation, image enhancement, contrast stretching, image thresholding, edge detection are implemented using XSG. Before going to implement algorithms, Gateway In and Gateway Out blocks should be connected in between pre-processing and post-processing blocks. These blocks act as input and output to the Xilinx portion of Simulink design.

4.1 Algorithm for Image Negative

The algorithm for gray scale image negative is given in Fig 5. Here, addsub block simply subtracts the constant 255 from pixel value to generate the negation of an input image. The output for this algorithm can be observed in Figure 6(b).
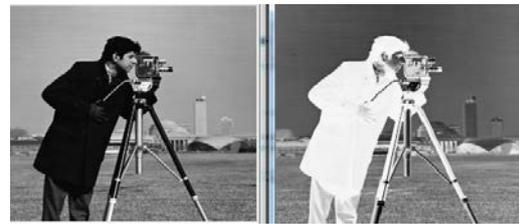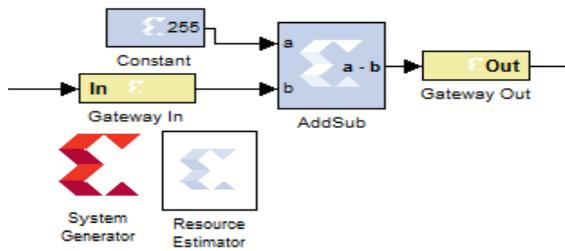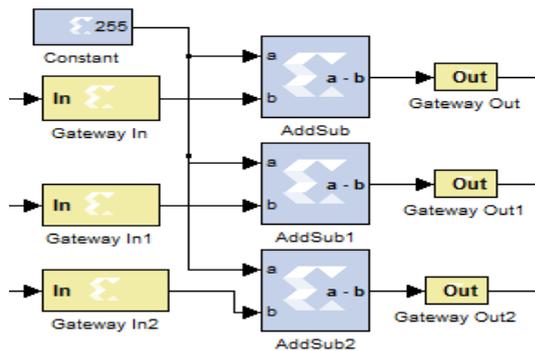

(a) Input Image   (b) Output Image

Figure 5: Algorithm for Gray Scale Image Negative   **Figure 6: Result for Gray Scale Image Negative**

For color images the algorithm is similar but implemented for three multidimensional R,G,B signals.

(a) Input Image   (b) Output Image

Figure 7: Algorithm for Color Image Negative   **Figure 8: Result for Color Image Negative**

4.2 Algorithm for Image Edge Detection

Edge is a basic feature of an image. Edge detection produces an edge map that contains important information about the image. Edge detection is simply the masking operation with suitable filter mask. 5x5 Filter Mask provides coefficients for an Edge. 5x5 filters comprises of 5 n-tap MAC FIR filters, each MAC filter has separate architecture which consists of counter, addressable shift register, coefficient ROM, capture register, MAC engine. The absolute value of the FIR filters is computed and the data is narrowed to 8-bits. Sobel Y, Sobel X, Sobel X-Y, Smooth, Blur, Sharpen, Gaussian, or Identity filtering can be accessed. For Filtering operation the delay is created by 5x5 filter block. To avoid this error the system generator has to be clocked 5 times faster than the normal clock. Internal architecture for 5x5 filters and its internal diagram is given in Figure 9 and Fig. 10
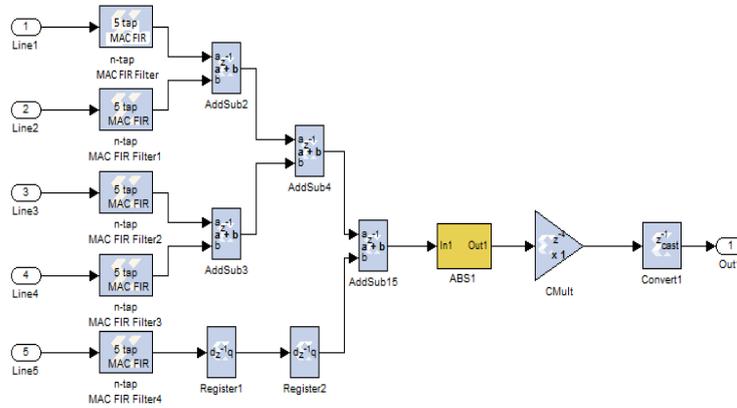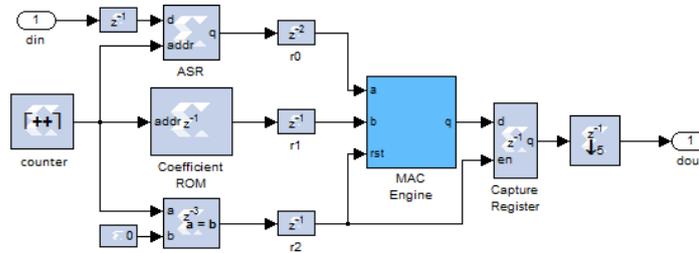
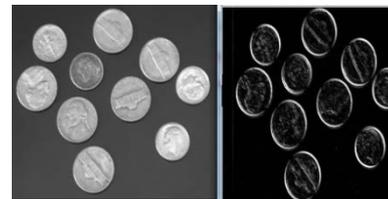**Figure 9: Internal Architecture for 5x5 Filter**

Figure 10: Internal Architecture for 5 tap MAC Filter

Algorithm and result for gray scale image edge detection is shown in Fig 11 and Fig 12(b)
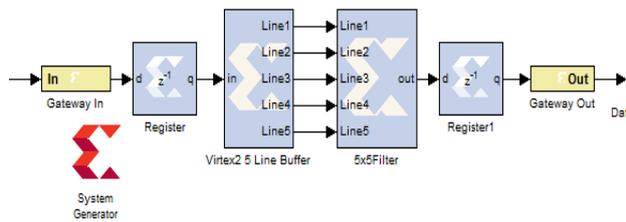
(a) Input Image   (b) Output Image

Figure 11: Algorithm for Gray Scale Image Edge Detection   **Figure 12: Result for Gray Scale Image Edge Detection**

Algorithm for color image edge detection is given below. It is similar to Gray scale image edge detection but is implemented for R, G, B signals.
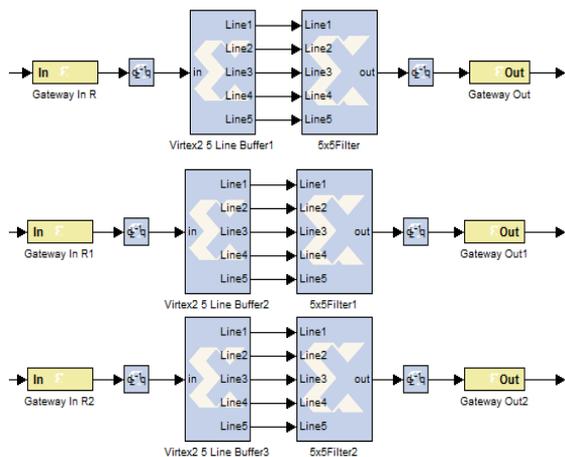
(a) Input Image   (b) Output Image

Figure 13: Algorithm for Color Image Edge Detection   **Figure 14: Result for Color Image Edge Detection**

4.3   Algorithm for Image Thresholding

Thresholding an image is the method of making all pixels are white above a certain threshold value while others black. For implementing the algorithm a suitable constant is taken e.g. 55, a Mux is used for replacing the thresholds by white. Mathematically it is expressed in equation (1) and (2).

$$\text{White (255) = old value > 55} \tag{1}$$

$$\text{Black (0) = old value < 55} \tag{2}$$

A binary decision can be carried out for each pixel using the rule

$$F(x, y) = \begin{cases} 1 & \text{if } f(x, y) > T_{xy} \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

Where T_{xy} is the threshold assigned to location (x, y) in the image

Algorithm and result for image thresholding can be observed in Figure 15 and 16(b)
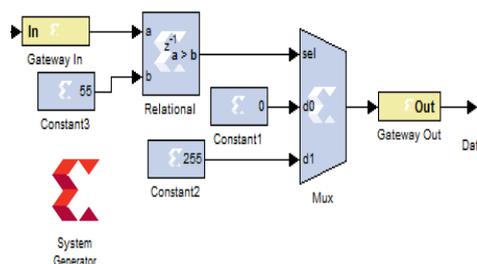
Figure 15: Algorithm for Image Thresholding

(a) Input Image   (b) Output Image

Figure 16: Result for Image Thresholding

4.4   Algorithm for Contrast Stretching

Contrast stretching means changing the contrast or luminescence values of an image. It can be achieved by using constant multiplier and add sub blocks and basic algorithm is given below. Mathematically it is expressed in equation (4).

$$\text{New value} = \{(\text{old value}-240)*3\} + 238 \tag{4}$$

Algorithm and result for contrast stretching can be observed in Figure 17 and 18(b)

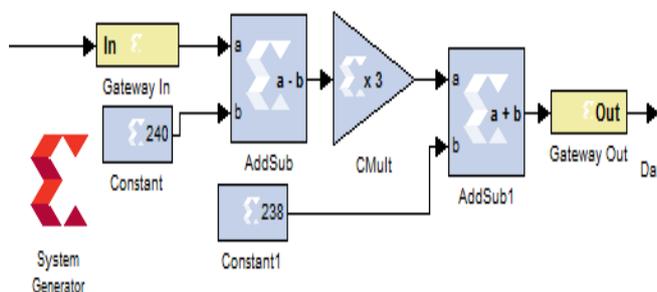**Figure 17: Algorithm for Contrast Stretching**

(a) Input Image   (b) Output Image

Figure 18: Result for Contrast Stretching

4.5   Algorithm for Image Enhancement

Images can be enhanced by adding some constant values to it. The basic algorithm and its result were shown in Figure 19 and Figure 20(b).

(a) Input Image   (b) Output Image
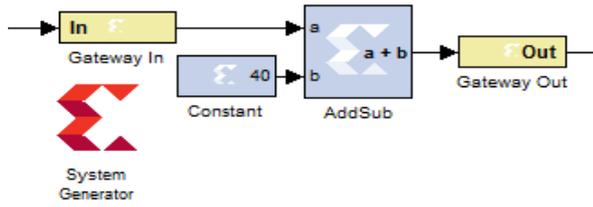
Figure 19: Algorithm for Image Enhancement   **Figure 20: Result for Image Enhancement**

4.6   Implementation of Image Processing Lab

In this section, grouping of various image processing algorithms is done simply by using multiplexer. It gives the output for any of these algorithms connected to (d0-d4) lines of multiplexer by giving some constant value as shown in Figure 21. For example 2 is given as constant, it gives the image thresholding output shown in Figure 22(b).

Figure 21: Implementation of Image Processing Lab

(a) Input Image   (b) Output Image

Figure 22: Result for Image Processing Lab for Constant 2

5   Implementation using Xilinx

Before processing the images on FPGA, HDL code should be generated by using system generator token. The system generator token is used to generate the code. Every Simulink model must contain system generator token to generate the code for algorithms designed in Simulink. Once a system generator token is added to a model, it is easy to generate code for the designed model. To generate 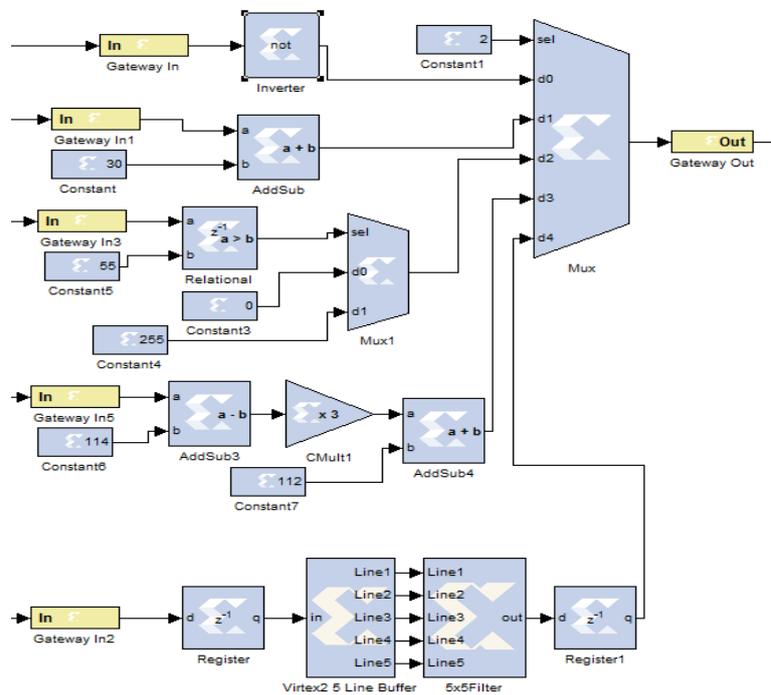the code, one should follow the procedure. First step in the procedure is opening of system generator token and selection of, compilation type which specifies the type of compilation that should be produced when the code generator is invoked. HDL net list compilation type is used in every model. Part, which the specific parts to be used on FPGA. Virtex5 xc5vlx100t-3ff1136 FPGA is used in every model. Synthesis tool specifies the tool used to synthesize the design. XST tool was used as a synthesis tool for the models described in the section. Hardware Description Language, Specifies the HDL language used for compilation of the design. In every model, VHDL is set for hardware description language. Create test bench, this intimate's system generator to create a HDL test bench.

 Second step in the procedure is clocking tab. The parameters for the clocking tab are as follows.

- *FPGA Clock period; defines the period in   nanoseconds of the system clock.*
- *Clock pin location; defines the pin location for the hardware clock.*

After setting the above parameters, press generate button in Xilinx system generator token. Then it automatically generates the code in specified path in Target directory. Open ISE navigator window and open the code generated files and simulate the code. Automatically it generates the RTL and Technology schematic diagrams, synthesis reports and timing diagrams. RTL schematic and Timing diagrams for image processing lab are shown in Figure 23 and Figure 24.
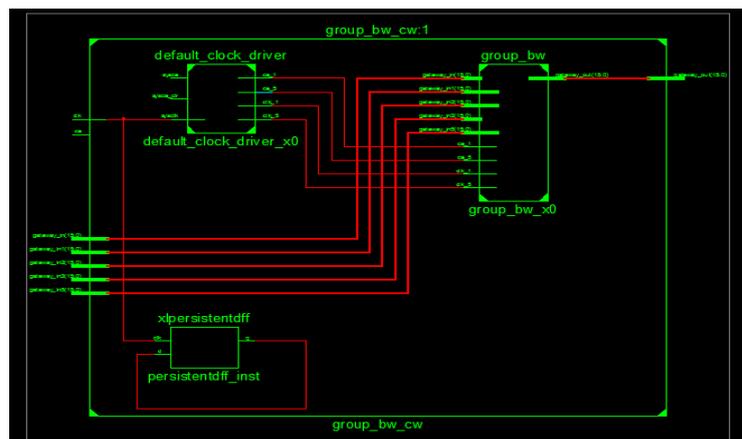

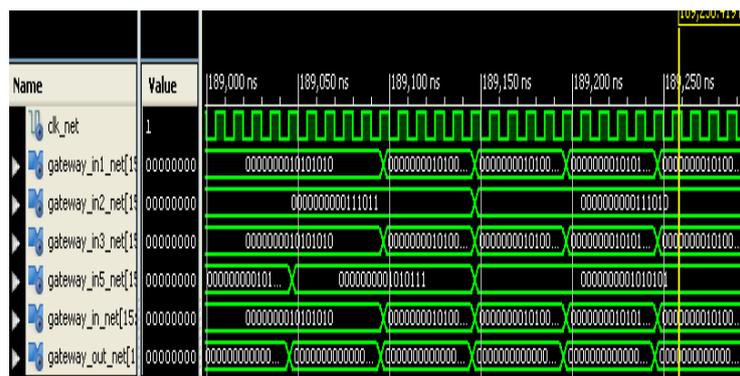**Figure 23: RTL Schematic for Image Processing Lab**

Figure 24: Timing Diagram for Image Processing Lab

6   Conclusion

Xilinx system generator is very helpful tool for software and hardware image processing tasks. It provides means to do hardware implementation of image processing algorithms with minimum resource and minimum delay. Thus, Matlab, Simulink and Xilinx system generator tools are extremely important in today's world as technology grows rapidly, that is, large number of pixels is being crammed into ever-smaller spaces. It provides easy hardware implementation.

7   Future Scope

The same concept can be extended to various fields like video processing, computer graphics, animations etc. The paper describes simulation part and suitable results are observed. The same results can also be verified on FPGA by dumping the code generated by system generator token.

REFERENCES

[1]. D.Selvathi, J.Dharani,"*Realization of Beamlet Transform Edge Detection Algorithm Using FPGA*", International Conference on Signal Processing, Image Processing and Pattern Recognition (ICSIPR), PP.1-5, 2013.

[2]. Jinbo Wu, Zhouping Yin, and Youlun Xiong, "*The Fast Multilevel Fuzzy Edge Detection of Blurry Images*" IEEE Signal Processing Letters, vol.14, No.5, may 2007

[3]. C.Sujatha, D.selvathi,"*Hardware Implementation of image edge detection using XSG*", Asian Journal on Scientific Research, 2014.

[4]. Ravi.s, Abdul Rahim,"*FPGA based design and implementation of image edge detection using XSG*", international journal of engineering trends and technology (IJETT), vol.4, issue.10, PP.4657-4660, oct 2013.

[5]. Obili ramesh, P.V.Krishna Mohan Gupta, "*A Real Time Hardware And Software Co-Simulation of Edge Detecetion For Image Processing System*", International Journal on Engineering and Research Trends(IJERT), vol.2, issue.8, PP.1695-1701, aug-2013.

[6]. Simulink user's guide, 2010.

[7]. "*Hardware software co-simulation for traffic load computation using matlab and simulink model blockset*", International Journal of Computational Science and Information Technology (IJCSITY), Vol.1, IssueNo.2, May 2013.

[8]. "*FPGA Based Design and Implementation of Image Architecture using XILINX System   Generator* ", IJCAE, Vol. No. 3 Issue 1, July 2012, pp132- 138.

[9]. "*Accelerated Image Processing on FPGAs*", IEEE Transactions on Image Processing , Vol. 12, issue 12, Dec. 2003,pp1543-1551.

[10]. "*Architecture for filtering images using Xilinx system generator*", INTERNATIONAL   JOURNAL of MATHEMATICS AND COMPUTERS IN SIMULATION, Issue 2, Volume 1, 2007,pp 101-106.

[11]. "*Performance Efficient FPGA Implementation of 2D image filtering using XSG*", IJECT Vol. 3,   Issue 4, Oct - Dec 2012,pp 374-377.

[12]. Matlab hdl coder, http:// www.mathworks.in/products/hdl-coder.

[13]. Xilinx System Generator User's Guide, http://www.Xilinx.com/support/sw_ manuals/sysgen_user.pdf.

[14]. Digital Image Processing text book by Rafael C.Gonzalez, Richard

Top-down Spatial Attention for Visual Search: Novelty Detection-Tracking using Spatial Memory with a Mobile Robot

Nevrez Imamoglu[1], Enrique Dorronzoro[2],[3], Masashi Sekine[1], Kahori Kita[3] and Wenwei Yu[3]
[1]*Graduate School of Engineering, Chiba University, Japan,* [2]*Department of Electronic Technology, University of Seville, Spain,* [3]*Center for Frontier Medical Engineering, Chiba University, Japan;*
nevrez.imamoglu@chiba-u.jp; enriquedz@dte.us.es; sekine@office.chiba-u.jp; kkita@chiba-u.jp;
yuwill@faculty.chiba-u.jp

ABSTRACT

Assistive robotics technologies have been growing impact on at-home monitoring services to support daily life. One of the main research fields is to develop an autonomous mobile robot with the tasks detection, tracking, observation and analysis of the subject of interest in the indoor environment. The main challenges in such daily monitoring application, thus in visual search, are that the robot should track the subject successfully in several severe varying conditions. Recent color and depth image based visual search methods can help to handle part of the problems, such as changing illumination, occlusion, and etc. but these methods generally use large amount of training data by checking the whole scene with high redundancy to find the region of interest. Therefore, inspired by the idea that spatial memory can reveal novelty regions for finding the attention points as in Human Visual System (HVS), we proposed a simple and novel algorithm that integrates Kinect and Lidar(Light Detection And Ranging) sensor data to detect and track novelties using the environment map of the robot as a top-down approach without the necessity of large amount of training data. Then, novelty detection and tracking is achieved based on space based saliency map representing the novelty on the scene. Experimental results demonstrated that the proposed visual attention based scene analysis can handle various conditions stated and achieve high accuracy of novelty detection and tracking.

Keywords: Novelty detection and tracking, Visual attention computational models, Space based spatial saliency, Robotics, At-home monitoring.

1   Introduction

The significance and necessity of assistive technologies in medical and service industry have been growing as a research field for daily life [1-8]. Researchers should introduce projects to

improve the people's or society's quality of life; especially, for the people who are in demand such as elderly or people with disabilities. Recently, with the increase in population, it is obvious that the work load of service and medical industry such as the clinics and hospitals increasing rapidly that can decrease the efficiency and quality of the service to the people in demand [1-3]. As an emerging topic in the assistive intelligent systems field, in our previous work [3, 9, 10], we have been working on developing at-home bio-monitoring mobile robot system for motor function impaired persons (MIPs) with the tasks tracking, behaviour detection and analysis, and etc. (Figure 1), which have been becoming an important concept in medical field for those who have difficulties leaving their house such as elderly people or MIPs [3, 9, 10]. These systems can benefit by providing more convenient and comfortable way of health care or daily support for the patients and reducing the workload of

therapists that can lead to time and economy efficiency concerning the quality of life [3, 9, 10]. In addition, they can provide seamless safety conditions for the people at-home by enabling full observation of their behaviors in any condition, in which smart camera systems or wearable sensor may have problems such as blind spots or lack of data for analysis [3, 9, 10].
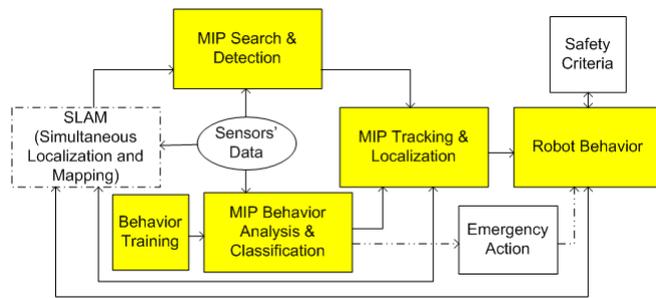

Figure 1: Necessary Modules for At-Home Monitoring Autonomous Mobile Robot System for Mips Observation (Highlighted Modules are the Topics of Interest in our Project)
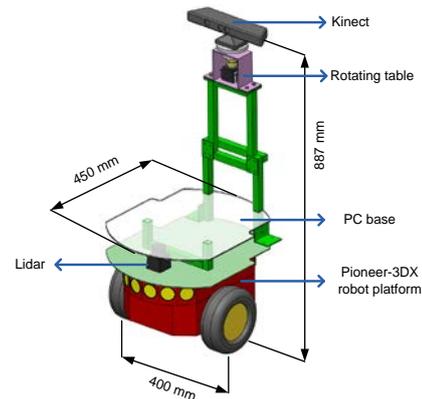
Figure 2: Bio-monitoring Mobile Robot Platform with Sensors to Provide the Tasks Defined in Figure 1

The required modules for an autonomous mobile robot system with monitoring skills are given in Figure 1 where the highlighted modules are the current research interests such as detection, tracking, subject's activity recognition, and analysis of walking pattern using Kinect RGB-D camera sensor. For such an at-home monitoring mobile robot (Figure 2), if the visual tracking module of the system is not robust enough, the whole system will be ineffective regardless of how good or accurate activity recognition or human model the system has. It is obvious that visual search for detection and tracking of subject of interest is the main core of at-home monitoring mobile robot system: it starts or supports all the at-home monitoring tasks, including subject localization, measurement, activity recognition, and robot control.

In this work, to achieve robust detection and tracking of the novelties on the scene with low complexity algorithms and high novelty detection accuracy based on range sensors, we proposed a fast and novel algorithm by integrating 2D (laser range finder: Lidar) and 3D (Kinect) sensor information. The proposed algorithm detects and tracks the points with the highest novelty on the scene by using the robot environment memory as a top-down approach inspired by the concept of spatial memory and visual attention mechanism in human visual system during visual search. So, we achieved to detect and track subject on the scene as a novelty of the environment extracted by using the proposed space-based saliency algorithm.

The paper is organized as follows. Section 2 demonstrates the related work as background information along with the idea and contribution of this work. Section 3 describes the proposed space-based saliency model based on the spatial memory, experimental results and discussions were given in Section 4, and finally, concluding remarks were stated in Section 5.

2   Background and Contribution of the Study

In order to achieve robust observation of human subjects at home, it is very critical to achieve effective visual search for the robot with subject detection and tracking tasks. The main challenges in such daily monitoring application, thus in visual search, are that the robot should detect and track the

subject successfully in several severe conditions, such as: (i) low quality data with distortion or noise, (ii) partial occlusion, (iii) highly changing illumination conditions, and etc.

2.1 Related Works in Literature

2.1.1 Background on Color Image based Detection and Tracking

Regarding the detection and tracking, many algorithms have been developed using color image sensors to handle these problems [11-13]. There are several ways to utilize detection algorithms; using interest point detectors, segmentation models, background and/or foreground modeling, and using supervised classifiers with extracted features from the images [13]. For example, Gupta *et al.* [14] took advantage of SURF [15] descriptors, inspired by SIFT [16] method, for matching features from a reference subject for detection and tracking of person by a mobile robot. However, SIFT, SURF, or similar interest point based algorithms on color camera images depend on texture details, enough image resolution, and scene quality [17]. Therefore, these models are sensitive to the high illumination changes, low texture details, and low resolution images [17]. Nguyen *et al.* [18] proposed a method to use interest points to find the search points for the objects (e.g. human), then using their proposed texture descriptor non-redundant local binary patterns for matching by a Bayesian approach to find the likelihood [18]. But, it still has the disadvantages of interest point detectors, and it utilizes window scanning on the image with several scales and positions, which can be time consuming for real-time applications.

Color histogram information can also be utilized to detect and track the region of interests on images. Ning *et al.* [19] proposed the Scale and Orientation Adaptive Mean Shift Tracking (SOAMST) algorithm, which also offers invariance to rotations and scale, and it also updates search region for each frame. Moreover, it can accomplish tracking even under low resolution and low quality image conditions so it can be a good candidate for indoor real-time robot based subject tracking applications [19]. On the other hand, it also has problems with high illumination changes, and the tracking may fail if the background region has similar representation to reference frame.

In sum, there are many powerful tracking algorithms using color images as expressed work or more in literature [14-19]. However, in general, for at-home daily observation, they can be easily affected by the image data or environmental conditions such as low resolution image, low texture quality of the image, and high illumination changes, subject pose or gait differences, and etc. Therefore, there are also applications to incorporate color image data with other sensory information for detection and tracking applications.

2.1.2 Background on Multi-Sensor Approaches for Detection and Tracking

With the advancement of sensor devices in imaging technologies and new algorithms in computer vision applications, different sensory type representations of the scene can be provided such as depth or disparity images, 3D point clouds, thermal images, and etc [12, 20-23]. For instance, Talha and Stolkin [20] proposed an adaptive system that utilize particle filter based subject tracking by fusing thermal and visible spectra camera. Using thermal cameras along with color images, it is possible detect and track human subjects accurately; however, thermal camera systems generally has low resolution field of vision, and high-quality and high-resolution thermal imaging solutions are too expensive to be considered in at-home bio-monitoring robot projects.

On the other hand, range sensor applications for range imaging and 3D point cloud representations, are more appropriate choice considering the cost and performance. Therefore, applications have

been developed for detection and tracking by using color and depth data integration [21-23]. For example, Garcia *et al*. [21] proposed particle filter model by representing the particle states with 3D world coordinates and particle features with depth gradients and polar representation of color averages from HSV color space Cartesian definition [21]. It is also possible to detect and track specific objects by using prior knowledge from disparity and depth images of the region of interest. In the study of [22], detection and tracking algorithm is proposed using stereo vision mounted on a mobile robot by using the head and shoulder shape information such height and width ratio from disparity [22]. In addition, Liu et al. [23] make use of 3D point cloud data obtained from Kinect to find the possible candidate points by selecting points with local height maxima compared to their surroundings [23]. Then, for learning the human appearance model with SVM classifier, color-height joint histogram features are extracted from various height and human head appearance (front, back, and etc.) [23]. Then, appearance model is used to track the human detected on the scene [23].

In sum, depth or 3D data fusion with color based features can greatly improve the detection and tracking performance of the system considering the illumination changes and object representation. However, the models that combines color and depth data can fail under high illumination conditions due to the change in color appearance model. And, most of the models requires large amount of training data depending on the detection task with supervised approaches. In addition, due to searching and checking all the interest points on the image, these models generally have redundancy. Especially for detection cases, all the images goes under high-level feature extraction, feature matching, or feature classification procedure, which is redundant and time consuming task on the visual search. Even though there are models to eliminate the redundant information as in [23], they lack of using the knowledge on the environment, which can greatly deal with the redundancy especially for the indoor applications regarding the at-home monitoring for daily support.

2.1.3 Background on Visual Attention (Saliency Map) based Detection and Tracking

Visual Attention (VA) mechanism is an important part of Human Visual System (HVS) by avoiding redundant data on the scene and popping out the significant information on the scene for the benefit of selective attention process [24, 25]. Visual attention mechanism works from two perspectives such as spatial attention (space based attention) and object base attention, where spatial attention is spatial features such as contrast and orientation but object based attention relies on the object structure such as shape [26-28]. By using these relations, VA mechanism can benefit scene analysis and visual search by using low-level features for task independent situations or prior information with decision making process if there is task or knowledge dependency during the visual search [24, 25]. Many computational models have been developed to mimic VA for the benefit of computer vision applications; however, most of the saliency models are based on the 2D color images [24, 25]. And, there are few applications to create saliency maps for specific detection and tracking approaches [29-31]. Zhang *et al*. [30] proposed a context aware saliency model for the key object discovery and tracking application by introducing spectral affinity to saliency computation and saliency based particle filter tracking [30]. Yang et al. [31] proposed a more task specific approach to create saliency maps for specific objects such as bicycle, car and person.

On the other hand, several approaches combined the 3D data integration to color based models for saliency computation [32-37]. As stated by Wang et al. [32], most of the 3D based models utilizes depth image from three perspectives [32]: i) depth weighting [33], ii) depth saliency [34], and iii) stereo vision models that compute and use disparity [35]. All the studies stated until now, they were generally space based approaches by comparing one position to another based on 2D color images or depth images for 3D models. There are few studies to utilize space and object based attention models

or 3D real world space by using all three *XYZ* dimensions to find attentional objects [36, 37]. For example, Garcia and Frintrop [37] proposed a model for attentional 3D Object Detection by using Kinect RGB-D scene, where they do combine clustered 3D data and color image saliency to create a 3D object saliency map [37]. However, these approaches do not tell the exact novelty of the scene since they do not use spatial memory of the scene. So, attention regions or attentional objects found from these algorithms can both includes novelty or previously existing irrelevant part of the scene. And, the result still redundant to give valuable information such as what is new in the scene or what is most likely to be to object or subject of interest for a monitoring system ,which is crucial for surveillance of people living alone and in demand. Therefore, there is a need for a better 3D approach for novelty detection to observe subject robustly in indoor environment.

2.2   Idea and contribution of the work

For the detection and tracking tasks, low level and high level features on 2D/3D space or objects, such as color contrast, orientation, shape, object appearance model, and etc., are certainly necessary for an attention mechanism to detect novelties on the scene either for bottom-up or top-down mechanisms. However, current models neglect knowledge or effect of the space based or spatial attention [26-28], regardless of object features during visual search, which can decrease redundancy on attention significantly. As stated and referenced by Chun and Wolfe [38], experience and memory also affect attention while the attention has its contribution to experience and memory [38, 39]. In addition, Oh and Kim [40] also demonstrated the influence of spatial memory on visual search. Most importantly, Chun and Wolfe [38] also states that, people tend to pay attention to the regions, items or objects with novelty on the scene [38, 41]. These physiological studies [26-28, 38-41] prove that novelty based on spatial information, which can decrease redundancy and guide human attention during visual search, can also be an important factor for attention process independent of object features. Therefore, inspired by the idea of these spatial memory and space-based attention on HVS, we proposed a computation model to employ spatial memory as the knowledge of occupied region in the environment. We tried to detect novelties, specifically person of interest for the at-home monitoring mobile robot, which does not belong to the part of occupied regions of environment. We implemented a detection and tracking algorithm based on finding the novelty on the scene with the idea of space-based attention.

On the other hand, similar to the saliency approaches that generally combine spatial information in 3D and 2D color information, the most similar studies other than saliency map are related to the semantic mapping or semantic representation [8, 42-44] of the scene which may also associate visual search and visual memory; however, most of the applications are based on the analysis of data for segmentation [42] or creating 3D semantic maps (semantic SLAM) [8, 43, 44] rather than novelty detection. For example, in the study [44], the system integrates different sensory information such that robot uses 2D global mapping of the environment for localization, and with Kinect, 3D local information is transformed into global frame to find out objects and their changes in the position if any during long-term observation in indoor [44]. So, these models do not detect novel regions temporarily exist in the scene such as persons.

In addition, regarding the novelty detection and tracking for monitoring, semantic scene analysis approaches do not consider the attention or saliency concept to define visual priority of the regions, or they do not search for novelty on the scenes, with high likelihood to be the subject of interest in a known environment, to decrease redundancy in visual search. Therefore, by integrating different sensory information as in [44], we proposed a novel space-based saliency approach to detect and

track novelties for at-home monitoring task to support people in demand such as elderly or MIPs who may be living alone. To the best of our knowledge, there have not been any models developed for novelty detection based on spatial memory of the real world scene, since most of the attention based approaches are either 2D model or depth data processing of the visual field for depth saliency computation. In addition, the algorithm can be integrated into many existing color or depth based state-of-the-art models for detection and tracking applications to decrease the search points on the scene. Because the algorithm can decrease the search points or regions by providing the novel regions as task-dependent high priority attention regions on the scene with the aid of spatial memory based attention model. In this work, we have tested the model with different walking patterns, illumination and environmental conditions on various recorded data. Experimental results demonstrated that proposed visual attention model can handle various conditions stated by having high accuracy of novelty detection and tracking.

3   Proposed NOVELTY Detection and Tracking

The idea is to generate a space based spatial saliency map, in which the detected novel point is tracked along the continuous frames. The flowchart of the proposed model is given in Figure 3. In the proposed model, space-based saliency map is computed by comparing the readings from Kinect sensor (Figure 2) mounted on the rotating Table with the prior information of occupied region on the environment memory from Lidar sensor (Figure 2) along with the depth and size heuristics of the candidate regions.
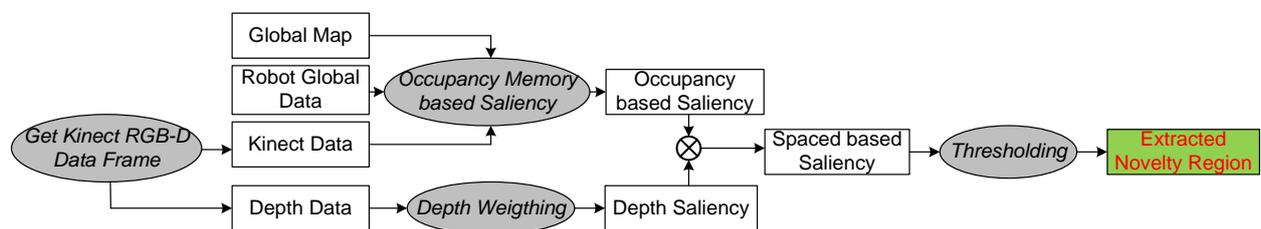

Figure 3: Flowchart of the proposed novelty detection and tracking algorithm

To be able to transform the local Kinect data to global environment map, first, the robot initially should build the map of room or house as a spatial memory of the environment. For robot handling and mapping tasks, we take advantage of Robot Operating System (ROS) [45], in which ROS incorporates with the ROSARIA package [46] to read the robot odometry data. And, prior to the main task of novelty detection and tracking, robot generates a 2D map of the environment with Gmapping package in ROS as described in [47, 48] by combining the odometry and Lidar (Figure 2) sensory information with 5cm grid resolution. Then, built map can be used for localization and robot path planning, as demonstrated in Figure 4(a), with ROS navigation stack that includes localization, costmap, global and local path planer. Robot pose is estimated by using the Adaptive Monte Carlo localization algorithm proposed in [49] from the global map generated with Gmapping [47, 48]. And, 2D occupancy grid map as the cost map [50] is employed to define the cost of grids on the map, which consists of obstacle information, static map, and dynamic sensory information for update process.
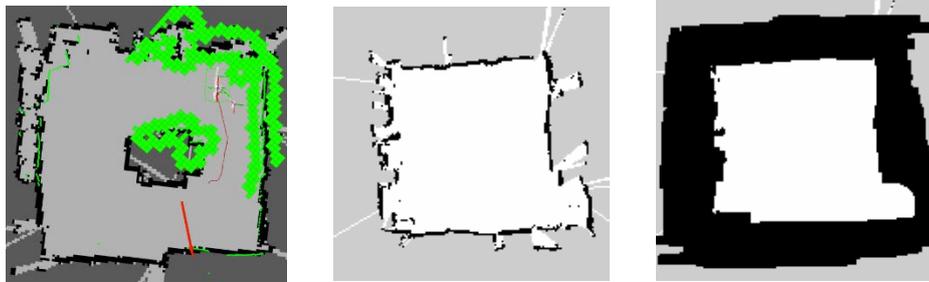
Figure 4: (a) ROS localization and path planning on the global map built by robot exploration, (b) global 2D Occupancy map (c) updated occupied memory map for the regions where novelties cannot exist (black pixels are occupied)

After having the global 2D map as the environment memory for the robot, we can use the global occupancy map and robot pose to find out the space based saliency. Initially, irrelevant places should be removed or added as occupied region to memory such as table plane or regions that subject cannot appear. Therefore, the locations such as tables or irrelevant regions with their surfaces on the global map (Figure 4(b)) were added to occupied regions of the memory (Figure 4(c)) where the subjects cannot be observed. Then, next step is to process Kinect 3D data for occupancy memory based saliency detection (Figure 3). First, we obtain the Kinect 3D data, and remove the floor and ceiling from the observation to reduce the irrelevant points for the saliency calculation by using height threshold values on 3D points (Figure 5). Also, since the Kinect data is 3D and the global map is 2D without any height values, Kinect data is reduced to 2D dimension by avoiding the height information. And, the occupancy memory based saliency (see Figure 3) of the points on filtered Kinect 3D data can be defined as:

$$\mathbf{S}_o \left(\mathbf{X}_k^l \middle| \mathbf{O}_m^g, p_r^g, \theta_r^g, \alpha_k^g, \mathbf{X}_k^l \right) \tag{1}$$

where $\mathbf{S}_o$ is the occupancy based saliency values of the observed points on Kinect local data given the local and global information such as $\mathbf{O}_m^g, p_r^g, \theta_r^g, \alpha_k^g$, and $\mathbf{X}_k^l$. In Equation (1), $\mathbf{X}_k^l$ is the local kinect data with horizontal and depth distances as the Kinect center is the origin, and depth (*x*) values of the locally detected occupied regions on the scene, $\mathbf{O}_m^g$ is the global occupancy map, p_r^g is the robot global position on 2D axis, θ_r^g is the global pose of the robot on $\mathbf{O}_m^g$, α_k^g is the Kinect pose differs from robot pose due to the rotating table as shown in Figure 2.
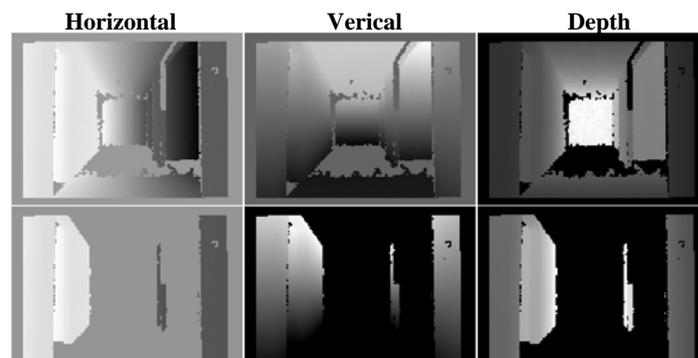

Figure 5: (top row) Kinect 3D data (bottom row) Kinect 3D observation after removal of floor and ceiling

To be able to calculate the occupancy based saliency in Equation (1), Kinect local data should be transformed into global values (Figure 6) regarding the global 2D map to be able to find out novelty points as in Equation (2) [51]:

$$\begin{bmatrix} K_{Gx} \\ K_{Gy} \end{bmatrix} = \mathbf{A} \times \begin{bmatrix} K_{Lx} \\ K_{Ly} \end{bmatrix} + \begin{bmatrix} T_{Gx} \\ T_{Gy} \end{bmatrix} \tag{2}$$

$$\mathbf{A} = \begin{bmatrix} s_x \cos(\alpha) & -s_y \sin(\alpha) \\ s_x \sin(\alpha) & s_y \cos(\alpha) \end{bmatrix} \tag{3}$$

$$\begin{aligned} T_{Gx} &= R_{Gx} - d\cos(\theta) \\ T_{Gy} &= R_{Gy} - d\sin(\theta) \end{aligned} \tag{4}$$

where K_{Gx} and K_{Gy} are the transformed Kinect global x and y positions on global 2D map, K_{Lx} (depth data in Fig.6) and K_{Ly} (horizontal data in Figure 5) are the local Kinect data on the relevant x and y axis. K_{Lx} and K_{Ly} is resolution is adjusted to 5cm resolution to match the 2D map resolution since each pixel representation 2D map image corresponds to 5cm grid regions. α is the Kinect pose on global map calculated by the robot pose and rotating table angle. $\mathbf{A}$ is the transformation matrix given in Equation (3) [51], in which s_x and s_y are the scaling coefficients on x and y axis. T_{Gx} and T_{Gy} are the translation values due to the difference between robot center and Kinect position on the robot (Figure 2). Translation values are defined as in Equation (4) where R_{Gx} and R_{Gy} are the robot position in global 2D map, and d is the distance of Kinect position to the robot center, and θ is the robot pose on global map.
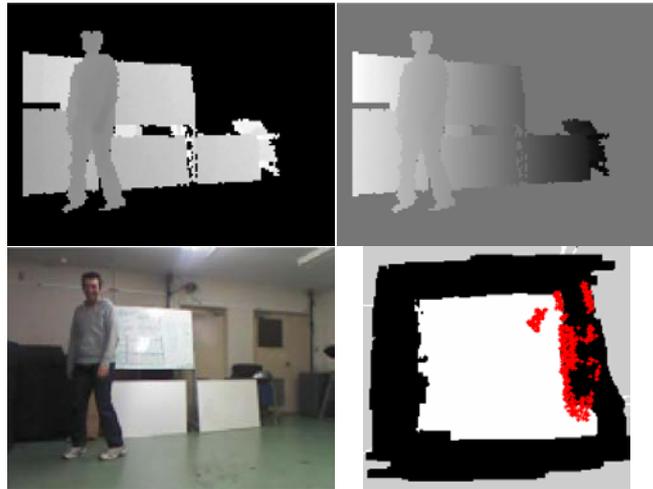
Figure 6: (top row) Kinect 3D local data depth and horizontal images respectively (bottom row) Color image of Kinect 3D observation, and $K_G(x,y)$ transformed points are shown on 2D map with red marks

For each transformed Kinect global points, the occupancy based saliency can be defined by find the distances of the global Kinect points to closest occupied regions on the global map as in Equation (5) with the given global transformed Kinect points ($K_G(x,y)$) and the environment memory ($\mathbf{O}_m^g$).

$$\mathbf{S}_o\left(K_{Gx}, K_{Gy}\right) = \arg\min\left(K_G(x,y), \mathbf{O}_m^g\right) \tag{5}$$

$$\textit{foreach } K_G(x,y)$$

$$\mathbf{S}_o\left(K_{Gx}, K_{Gy}\right) = \min\left(\text{norm}_{l^2}\left(\mathbf{O}_m^g, K_G(x,y)\right)\right) \qquad (6)$$

$$\textit{end}$$

where $S_o(K_{Gx}, K_{Gy})$ is the occupancy based saliency of the given $K_G(\underline{x},y)$ point with height h based on the environment knowledge $\mathbf{O}_m^g$ as in Equation (5). The definition of *argmin* function to calculate the saliency can be expressed in Equation (6), where the comparison of the Kinect points is done by l^2-*norm* of the vectors defined with the each Kinect point to the each map point; in other way, Euclidean distance of the Kinect points to the each occupied region in spatial memory.

Then, next step is to define the depth saliency to give more priority to the closer novelty regions by increasing the occupancy based saliency value of each $K_G(\underline{x},y)$. This operation is done by defining depth weights as in Equation (7).

$$\mathbf{S}_D = \frac{1}{\text{N}\left(\mathbf{X}_k^{depth}\right)+1} \qquad (7)$$

$$S_S\left(K_{Gx}, K_{Gy}\right) = S_D\left(K_{Gx}, K_{Gy}\right) \times S_o\left(K_{Gx}, K_{Gy}\right) \qquad (8)$$

$$S_S\left(K_{Gx}, K_{Gy}\right) = \begin{cases} S_S\left(K_{Gx}, K_{Gy}\right) & \textit{if } S_S\left(K_{Gx}, K_{Gy}\right) > T \\ 0 & \textit{otherwise} \end{cases} \qquad (9)$$

where $\mathbf{S}_D$ is the depth saliency map, which is used for weighting the occupancy based saliency map S_o, and N(.) is the normalization function for local depth data $\mathbf{X}_k^{depth}$ of Kinect sensor. For the calculated $\mathbf{S}_D$ and S_o, space-based saliency map, $\mathbf{S}_S$, can be calculated by Equation (8), where the saliency values less than a given threshold as in Equation (9) to remove irrelevant regions on extracted salient regions.
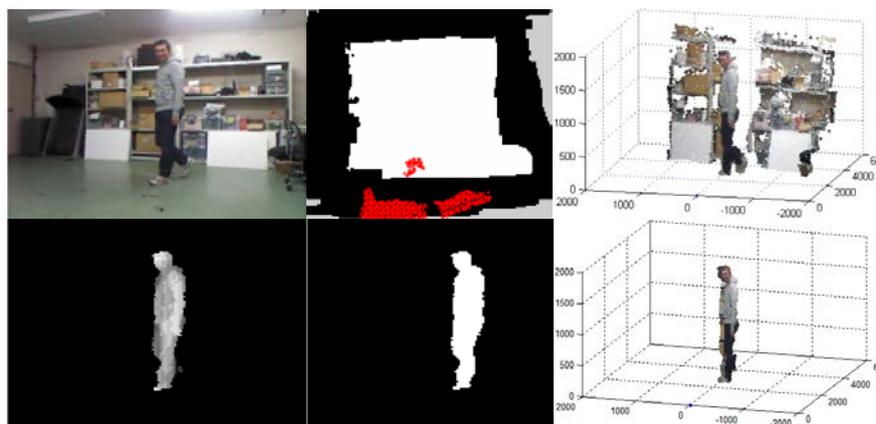

Figure 7: (top row) (a) Color image of Kinect 3D observation, (b) KG(x,y) transformed points are shown on 2D map with red marks, (c) representation of the Kinect data in 3D (bottom row) (a) space based saliency maps with height information of the bed, (b) extracted novelty region from space based saliency map, (c) space based saliency region in 3D

Since our study is to develop algorithms for at-home bio-monitoring, our priority is based on the large subjects (human), where small salient regions on the saliency map can be removed with a defined threshold based on 3D real data clustered regions or 2D image regions. Finally, novelty region extraction is simply done by segmenting all extracted regions, and finding their size on 2D image

representation, then the region with the largest size is selected as the novelty region to be tracked as in Figure 7.

In sum, using the prior information, the redundancy during visual search was decreased to find out the novelty of the scene, and then, some simple heuristics such as the closer depths and size of the candidate regions were used for the primary attention region decision as the tracking region on the scene, which are also consistent with the behaviour of HVS.

4   Experimental Results and Discussion

Before going through an extensive analysis of the algorithm, we started our test with a small number of test data from discrete data of some specific selected cases, in which robot were placed on the map randomly as Subject 1 (S1) was moving on the scene or a scene without any subject. The spatial memory for this test is the global map obtained in a hall of the building (Figure 8).
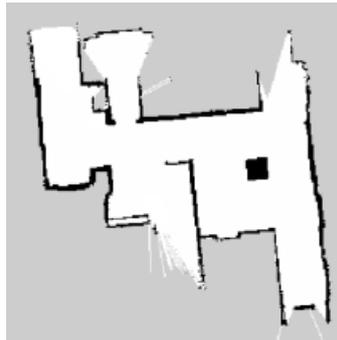

Figure 8: Occupancy map layout for the initial test with Subject 1

The dataset includes 55 frames of data including RGB color image, 3D Kinect data, robot pose, and Kinect pose. Among the 55 images, 22 data is without the subject as expressing the case that the scene does not have any novelty to be detected. And, 23 data includes the subject with different conditions of the subject and environment such as: (i) subject is fully visible, (ii) subject is partially visible (partial occlusion), (iii) subject is in various distance to the robot, and (iv) scenes are under various illumination conditions. Then, using the proposed algorithm in previous section, we analyzed each frame to detect novelties on Dataset-1. In Figure 9, some sample color images of Kinect data are given for the stated conditions with their space based saliency and extracted novelty region outputs.

It can be seen that when there is no novelty (subject) in the scene as in Figure 9(a), space based saliency and extracted novelty regions yields nothing as output. On the other hand, subject can be extracted as novelty from the space based saliency maps (Figure 9(f-h)) as in Figure 9(j-l) by using the spatial memory (global map in Figure 8) robot global data, and Kinect data. With the initial analysis, we tested detection performance in a simple environment with different situations like partial occlusion or partially visible cases (Figure 9(c-d)).

In Table 1, detection results are given with defined data (Dataset-1). For the 23 data with novelty (subject visible), number of true positives (TP) are 23 without any detection error, and the algorithm yielded no false positives (FP), which means that it did not extracted any non-novelty region as novel area instead of the subject. Also, for the 22 data without any novelty, number of true negative (Figure 9(a, e, i) cases) is 22 without any error, and the algorithm did not stated subject as as a part of spatial memory or non-novelty region.

Figure 9: (a) sample image without subject or novelty, (b-d) sample images with subject or novelty representing different conditions such as subject visibility or occlusion and illumination changes, (e-h) space-based saliency maps of the scenes in (a) to (d) respectively, (i-l) extracted novelty regions from saliency maps in (e) to (h)

Table 1: Detection results of the algorithm for Dataset-1

	Conditions	# frames	Accuracy
Novelty Exist (S1) & Without Novelty (55 frames)	TP	23	100%
	FP	0	100%
	TN	22	100%
	FN	0	100%
	Overall correct detections	55	100%

In addition, we also analyzed how the proposed algorithm can decrease redundant data for future processing based on novelty region. Our data consists of 120x160 pixel resolution, so it means that the number of points in point cloud is 19200. However, with the initial floor and ceiling removal some of these points neglected. With this removal on Dataset-1, average number of points to be processed decrease to 10883 points, where around 43.32% of the points were removed as redundant data during the initial process. Then, we calculated the average number of points remained on Dataset-1 after proposed novelty detection algorithm applied. On Dataset-1 with novelty existence cases (23 frames as given in Table 1), the average number of interest points decrease to around 2283 points with the proposed novelty extraction, which means that 11.89% of the data requires attention, and 88.11% of the points are redundant in average regarding the Dataset-1.

Experiment on randomly selected data with specific cases gives promising results to go for the next step with a more realistic environment and continuous data, where the subject moves in the room randomly to test detection and tracking performance of the algorithm. In the following section, detection and tracking performance of the systems is tested with more complex conditions.

4.1 Experiments for comparison

The proposed algorithm is tested with several conditions and compared with selected tracking algorithms by using a different dataset (Dataset-2) including more subjects and scenarios. Since the main purpose of the project is to observe the elderly or people with impairment, detection and tracking model should be able to track persons with different walking patterns in various indoor

environmental conditions. For the purpose of testing various simulated conditions, first, we generated a test dataset that consists of 6 recordings from 2 subjects with various illumination conditions, color variations and different walking patterns such as normal walking, simulated impaired walking, and simulated elderly walking.

Table 2: Properties of Dataset-2

Data	#image frames	Subject color	Subject exist in all frames	Light Changes	No-light condition
S2IW	2220	Black	Yes	No	No
S2EW	1245	Green	Yes	No	No
S2NW	1180	Black	Yes	No	No
S3IW	2120	Red	Yes	No	No
S3EW	2000	Red	Yes	No	No
S3NW	2390	Gray	No (172 frames)	Yes (decreased)	Yes (dark room)
Total:	11155				

The conditions for each recording data are given in Table-2. First of all, it should be noted that each data includes walking patterns inside the room with changing viewing point or angle. In Table-2, for the datasets that includes 11155 frames in total, recording name abbreviations can be described as; (i) S2IW: Subject 2 and impaired walking, (ii) S2EW: Subject 3 and elderly walking, (iii) S2NW: Subject 3 and normal walking, (iv) S3IW: Subject 3 and impaired walking, (v) S3EW: Subject 3 and elderly walking, (vi) S3NW: Subject 3 and normal walking (Figure 10).

During recordings of the datasets, subjects were wearing different color clothes for each recording set such as red, green, black and gray colors are the subjects color appearances for color tracking models (Figure 10). Red and green were chosen to be the easy to track color while black and gray colors are selected to test the robustness of the color tracker with similar background or highly changing illumination during subject motion. Moreover, in five dataset (S2IW, S2EW, S2NW, S3IW, S3EW), the conditions are set static such that (i) subject exists in all frames, and (ii) same light conditions are used during recording. However, based on the position of the subject, illumination value around the subject will be different due to the distance to the light sources. Moreover, the distance between the subject and robot will affect the texture details too. Hence, there are still dynamical conditions that make tracking difficult for color image data. On the other hand, one of the dataset (S3NW) is prepared with changing environmental and subject appearance conditions. For example, for the S3NW dataset, three different illumination conditions are used; (i) all the lights are on, (ii) illumination is decreased by closing some lights, and (iii) dark room environment as the all lights are off. In addition, because the proposed algorithm is based on novelty detection, existence and non-existence of the novelty region (subject in our case) should be tested too. Hence, in S3NW, 172 frames of 2390 frames do not represent any novelty as the subject is not on the visual area of the sensor.

Figure 10: Sample color images for each dataset (a) S2IW: Subject 2 and impaired walking, (b) S2EW: Subject 2 and elderly walking, (c) S2NW: Subject 2 and normal walking, (d) S3IW: Subject 3 and impaired walking, (e) S3EW: Subject 3 and elderly walking, (f) S3NW: Subject 3 and normal walking

4.1.1 Dataset-2 Tests and Comparison Results

With the given dataset, color tracking, depth improved tracking and proposed novelty detection and tracking algorithm are tested and compared. For the color tracker, we have chosen one of recent algorithms in the field, which is Scale and Orientation Invariant Adaptive Mean-Shift Tracker (SOAMST) [19] that can handle low resolution and texture images (e.g. 120x160 pixel images as in our case). The algorithm is easy to integrate it into real-time applications which require low computational cost. In addition, SOAMST [19] is proven to be quite robust to scale, orientation and illumination changes [19] as long as the region of interest (RoI) exists and is trackable on each frame. However, if the illumination value changes a lot that can alter the color information on RoI, SOAMST may not achieve tracking robustly as also stated prior in the related literature review section.

Due to the fact that, Kinect sensor provides depth information locally given the robot position, to compare with the proposed model, we also integrated depth likelihood map to improve the accuracy of tracking similar to the study in [10]. Although depth likelihood map integration, depth improved SOAMST (DI-SOAMST), can improve the result of color tracking [19], it is still not robust enough with the motion of subject and the robot. Moreover, when tracking fails ones in SOAMST and DI-SOAMST, it is very difficult to recover since both algorithms use previous frame's tacking points as a reference to continue next tracking process. If the previous frame's reference tracking point is false or not the subject of interest, then it is highly probable that next process output for the tracking may fail too.

Table 3: Comparison of the Tracking Algorithms with the Dataset-2

Dataset-2	SOAMST [19]			DI-SOAMST [19, 10]			Proposed Model		
	Detection & Tracking			Detection & Tracking			Detection & Tracking		
	Success	Track Fail	Detect Fail	Success	Track Fail	Detect Fail	Success	Track Fail	Detect Fail
S2IW	1668	552	NA	2220	0	NA	2220	0	0
S2EW	968	307	NA	1245	0	NA	1235	0	5
S2NW	1042	138	NA	1180	0	NA	1180	0	0
S3IW	2120	0	NA	2120	0	NA	2120	0	0
S3EW	2000	0	NA	2000	0	NA	2000	0	0
S3NW	1633	585	NA (172)	1484	734	NA (172)	2390	0	0
Total	9431	1582	172	10249	734	172	11150	0	5
Accuracy	85.60% (Tracking only)			93.32% (Tracking Only)			99.96% (Detect & Track)		

In Table 3, the results of tracking tests are given for SOAMST, DI-SOAMST, and proposed novelty detection algorithm. It should be noted that SOAMST and DI-SOMAST tracking only models, where they cannot detect or recognize whether the novelty exist or not (subject in the scene or not). Therefore, their detection and tracking results were not count if the subject does not exist in the scene. And, their detection performance (Detect Fail in Table 3) are assigned as not available (NA). On the other hand, the proposed model is detection and tracking algorithm so tracking and detection failures are obtained from the tests.

First of all, from results of S2IW and S2NW, it can be seen from Table 3 that walking pattern may also affect tracking performance of the color based approaches. In these two cases, Subject 2 has the same color appearance model and the layout is the same, the only difference between S2IW and S2NW is the walking pattern (Table 2), where S2IW is simulated impaired walking (Figure 10(a)) and S3NW is the normal walking (Figure 10(c)). And, while S2NW has 83.70% accuracy by having more stable walking gait, S2IW has lower performance as 75.14% tracking accuracy due to the high motion and gait change of the subject that can affect appearance model. On the other hand, color based tracking by using SOAMST can yield very high accuracy independent of walking pattern if the subject appearance model is very distinctive from the environment as in S3IW and S3NW (Table 2) with 100% tracking accuracy (Table 3). And, changing illumination conditions can decrease the tracking performance down to the 73.63% as in S3NW case. In sum overall tracking only performance of the SOAMST is 85.60% from all cases. On the other hand, DI-SOAMST showed promising results by handling SOAMST failures in five cases (S2IW, S2EW, S2NW, S3IW, S3EW) by giving 100.00% tracking performance (Table 3) without any failure when the light condition of the room does not change as in S2IW, S2EW, S2NW, S3IW, S3EW data (Table 2). For the S3NW data, DI-SOAMST is also affected from the changing light conditions of the room (Table 2). DI-SOAMST also failed to track subject in many frames of S3NW data due to loss of appearance model of the subject on the scene with highly changing illumination conditions or dark room. In general, it improved the tracking performance of SOAMST from 85.60to 93.32% as given in Table 3.

By having problem with detection after subject disappearance and reappearance, or with changing illumination conditions, as a color based approaches, both SOAMST and DI-SOAMST is not good enough to handle all day (day and night) monitoring task considering subject detection and tracking. However, Table 3 demonstrates that proposed space based novelty detection and tracking algorithm is a good candidate and a reliable model for monitoring by having 99.96% detection and tracking

accuracy among 11150 frames of all cases. Also, detection and tracking output of a frame is not affected by the detection or tracking errors in previous frames, which is not the case for SOAMST, DI-SOAMST or many color based state of the art tracking approaches.

In sum, proposed space based saliency model for novelty detection and tracking task yielded promising results for at-home monitoring mobile robot project since all the datasets in indoor environment and their results have very high accuracy. Moreover, this algorithm can be extended or integrated to other algorithms easily if more complex approaches are necessary for indoor surveillance tasks such as multiple person tracking and identification, activity recognition, change detection, and etc.

5   Conclusion

In this study, we demonstrated that the idea of spatial working memory during visual search as in Human Visual Systems (HVS) can be used to calculate saliency map from real world data. Therefore, we proposed an algorithm to create space-based saliency map for novelty detection and tracking by fusing two different sensors, in which Lidar is used for 2D global mapping and localization, and Kinect is used for perceive the local 3D data of the scene. Then, using prior environment knowledge, we showed that it is possible to detect and track subjects as the novelty of the scene by paying attention to the region that is different from the spatial memory, in other words, global map. With this approach, an efficient and fast model is obtained for indoor mobile robot based tracking by having more than 99.00% detection and tracking accuracy in tested Datasets.

As a future work, this model can be improved by using 3D global mapping and localization or height information integration for more complex environments instead of 2D approach even though current approach is good enough to detect people for monitoring walking activities. Also, it can be integrated with semantic map algorithms to work on a more semantic level. Moreover, multiple person detection and identification algorithms in more complex situation can be handled faster by decreasing the redundancy of RGB-D based algorithms with the proposed novelty detection.

REFERENCES

[1]. Rahidi, P. and A. Mihailidis, *A survey on ambient-assisted living tools for older adults*. IEEE Journal of Biomedical and Health Informatics, 2013. **17**(3): p. 579-590.

[2]. Gross, H.-M., et al., *Progress in developing a socially assistive mobile home robot companion for the elderly with mild cognitive impairment*. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS), 2011. p. 2430-2437.

[3]. Nergui, M., et al., *Human gait behaviour classification on lower body triangular joint features*. IASTED Int. Conf. on Signal and Image Processing (SIP), 2012. p. 212-219.

[4]. Cesta, A., et al., *The ROBOCARE assistive home robot: Environment, features, and evaluation*, The ROBOCARE Technical Reports, 2004. RC-TR-0906-6.

[5]. Jayawardena, C., et al., *Design, implementation and field tests of a socially assistive robot for the elderly*. The Fourth IEEE RAS/EMBS Int. Conf. On Biomedical Robotics and Biomechatronics, 2012. p. 1837-1842.

[6]. Li, R. and M. A. Oskoei, H. Hu, *Towards ROS based multi-robot architecture for ambient assisted living*, IEEE Int. Conf. on Systems, Man, and Cybernetics, 2013. p. 3458-3463.

[7]. Simpson, R. C., et al., *NaVChair: An assistive wheelchair navigation system with automatic adaptation*, Assistive Technology and AI, Lecture Notes in Artificial Intelligence (LNAI), 1998. **1458**: p. 235-255.

[8]. Wei, Z., W. Chen, and J. Wang, *Semantic mapping for smart wheelchairs using RGB-D camera, Journal of Medical Imaging and Health Informatics*, 2013. **3**(1): p.94-100.

[9]. Myagmarbayar, N., et al., *Human Activity Recognition Using Body Contour Parameters Extracted from Depth Images*. Journal of Medical Imaging and Health Informatics, 2013. **3**(3): p. 455–461.

[10]. Imamoglu, N., et al., *An Improved Saliency for RGB-D Visual Tracking and Control Strategies for a Bio-monitoring Mobile Robot*. Communications in Computer and Information Science, 2013. **386**: p. 1-12.

[11]. Liu, H., S. Chen, and N. Kubota, *Intelligent video systems and analytics: A survey*. IEEE Trans. on Industrial Informatics, 2012. **8**(1): p. 49-60.

[12]. Luo, R. Cc. and C.-C. Chang, *Multisensor fusion and integration: A review on approaches and its applications in mechatronics*. IEEE Trans. on Industrial Informatics, 2013. **9**(3): p. 1222-1233.

[13]. Yilmaz, A., O. Javed, and M. Shah, *Object tracking: A survey*. ACM Computing Surveys, 2006. **38**(4) - Article 13: p. 1-45.

[14]. Gupta, A.M., et al., *An on-line visual human trcking algorithm using SURF-based dynamic object model*. IEEE Int. Conf. on Image Processing (ICIP), 2010. p. 3875-3879.

[15]. Bay, H., et al., *Speeded-up robust features (SURF)*. Computer Vision and Image Understanding, 2008. **110**: p. 346-359.

[16]. Lowe, D.G., *Distinctive image features from scale-invariant keypoints*. International Journal of Computer Vision (IJCV), 2004. **60**(2): p. 91-110.

[17]. Borji, A., et al., Adaptive object tracking by learning background context. IEEE Int. Conf. on Computer Vision and Pattern Recognition (CVPR), 2012. p. 23-30.

[18]. Nguyen, D.T., et al., *Object detection using non-redundant local binary patterns*. IEEE Int. Conf. on Image Processing (ICIP), 2010. p. 4609-4612.

[19]. Ning, J., et al., *Scale and orientation adaptive mean shift tracking*. IET Computer Vision, 2010. **6**(1): p. 52-61.

[20]. Talha, M. and R. Stolkin, *Particle filter tracking of camouflaged targets by adaptive fusion thermal and visible spectra camera data*. IEEE Sensors Journal, 2014. **14**(1): p. 159-166.

[21]. Garcia, G.M., et al., *Adaptive multi-cue 3D tracking of arbitrary objects*. Pattern Recognition, Lecture Notes in Computer Science (LNCS), 2012. **7476**: p. 357-366.

[22]. Jia, S., et al., *Robust human detecting and tracking using varying scale template matching*. IEEE Int. Conf. on Information and Automation, 2012. p. 25-30.

[23]. Liu, J., et al., *Real-time human detection and tracking in complex environmentss using single RGBD camera*. IEEE Int. Conf. on Image Processing (ICIP), 2013. p. 3088-3092.

[24]. Fang, Y., et al., *Saliency detection in the compressed domain for adaptive image retargeting*. IEEE Trans. on Image Processing (IEEE TIP), 2012. **21**(9): p. 3888-3901.

[25]. Imamoglu, N., W. Lin, and Y. Fang, *A saliency detection model using low-level features based on wavelet transform*. IEEE Trans. on MultiMedia, 2013. **15**(1): p. 96-105.

[26]. Fink, G.R., et al., *Space-based and object-based visual attention: shared and specific neural domains*. Brain, 1198. **120**: p. 2013-2028.

[27]. Logan, G.D., *The CODE theory of Visual Attention: An integration of space-based and object-based attention*. Psychological Review, 1996. **103**(4), p. 603-649.

[28]. Mozer, M.C. and S.P. Vecera, *Space-and object-based attention*. Neurobiology of Attention, Academic Press, 2005. p. 130-134.

[29]. Frintrop, S., et al., *A component based approach to visual person tracking from a mobile platform*. International Journal of Social Robotics, 2010. **2**(1), p. 53-62

[30]. Zhang, G., Z. Yuan, and N. Zheng, *Key object discovery and tracking based on context aware saliency*. International Journal of Advanced Robotic Systems, 2013. **10**(15): p. 1-12.

[31]. Yang, J. and M.-H. Yang, *Top-down visual saliency via joint CRF and dictionary learning*. IEEE Int. Conf. on Computer Vision and Pattern recognition (CVPR), 2012. p. 2296-2303.

[32]. Wang, J., et al., *A computational model of stereoscopic 3D visual saliency*. IEEE Trans. on Image Processing (IEEE TIP), 2013. **22**(6): p. 2151-2165.

[33]. Chamaret, C., et al., *Adaptive 3D rendering based on region-of-interest*. Proc. of SPIE, 2010. **7524**: p. 75240V.

[34]. Ouerhani, N. and H. Hugli, *Computing visual attention from scene depth*. IEEE 15th International Conf. on Pattern Recognition, 2000, **1**: p. 375 -378.

[35]. Fang, Y., et al., *Saliency detection for stereoscopic images*. IEEE Trans. on Image Processing (IEEE TIP), 2014. **23**(6): p. 2625-2635.

[36]. Begum, M., et al., *Object- and space-based visual attention: An integrated framework for autonomous robots*. IEEE/RSJ Inter. Conf. on Intelligent Robots and Systems (IROS), 2008. p. 301-306.

[37]. Garcia, G.M. and S. Frintrop, *A computational framework for attentional 3D object detection*. Proc. of the Annual Meeting of the Cognitive Science Society, 2013.

[38]. Chun, M.M. and J.M. Wolfe, *Visual Attention*. Handbook of Sensation and Perception (Chapter 9), Edited by E. B. Goldstein, Blackwell Publishing, 2005. p. 273-310.

[39]. Chun, M.M. and K. Nakayama, *On the functional role of implicit visual memory for the adaptive deployment of attention across scenes*. Visual Cognition, 2000. **7**: p. 65-81.

[40]. Oh, S.-H. and M.-S. Kim, *The role of spatial working memory in visual search efficiency*. Psychonomic Bulletin & Review, 2004. **11**(2): p. 275-281.

[41]. Johnston, W.A., et al., *Attention capture by novel stimuli*. Journal of Experimental Psychology, 1991. **119**(4): p. 397-411.

[42]. Zhou, Y., et al., *Region based high-level semantics extraction with CEDD*. 2nd IEEE Int. Conf. on Network Infrastructure and Digital Content, 2010. p. 404-408.

[43]. Nuchter, A. and J. Hertzberg, *Towards semantic maps for mobile robots*. Robotics and Autonomous Systems, 2008. **56**: p. 915-926.

[44]. Mason, J. and B. Marthi, *An object-based semantic world model for long-term change detection and semantic querying*, IEEE/RSJ Inter. Conf. on Intelligent Robots and Systems (IROS), 2012. p. 3851-3858.

[45]. *Robot Operating System (ROS)*, http://wiki.ros.org/

[46]. *ROSARIA*, http://wiki.ros.org/ROSARIA

[47]. Grisetti, G., C. Stachniss, and W. Burgard, *Improved techniques for grid mapping with rao-blackwellized particle filters*. IEEE Transactions on Robotics, 2007. **23**: p. 34-46.

[48]. *ROS-Gmapping*, http://wiki.ros.org/gmapping

[49]. Thrun, S., W. Burgard, and D. Fox, *Probabilistic robotics*. The MIT Press Cambridge, 2005.

[50]. *Costmap*, http://wiki.ros.org/costmap_2d

[51]. *Transformation Matrix*, http://mathforum.org/mathimages/index.php/Transformation_Matrix

AIVP ADVANCES IN
IMAGE AND VIDEO PROCESSING

An Efficient and Effective Framework for Eye Region Detection and Eye State Recognition

Cheng-Chieh Chiang

Department of Information Technology, Takming University of Science and Technology
Taipei, Taiwan
kevin@csie.ntnu.edu.tw

ABSTRACT

This paper proposes a framework to treat the eye information in a face image including the eye region detection and the eye state recognition. In order to make it possible to employ our approach to a real application in practice, our goal in this paper is to design not only a fast enough but also a high performance framework for the eye region detection and the eye state recognition. Our proposed framework mainly contains two parts: the first is to locate the eye regions in a face image, and the second is to recognize the states, either open or closed, of the eye regions. When a frame is captured from a video sequence, a face detection method is first performed to determine the positions of face regions. Next, face regions are converted into binary images and then we perform the horizontal and the vertical projections to locate the eye regions. Two visual features containing the intensity values of pixels and local binary pattern (LBP) are extracted from eye regions to classify the eye states with the support vector machine (SVM) approach. This paper also demonstrates a several experiments to present the efficiency and effectiveness of our proposed framework.

Keywords: Eye Region Detection, Eye State Recognition, Local Binary Pattern, Support Vector Machine.

1   Introduction

Computer vision has been widely applied to many domains in our life. For example, the eye tracking system [1][2] can automatically detect eye regions and then track the user gaze in order to understand what targets the users are looking at. Several important issues, including eye region detection/tracking/state recognition and eye as bioinformatics for identification, to treat eye images/videos in computer vision have been active in recent. This paper presents our effort to propose a framework that cannot only locate eye regions but also recognize the eye states in a video frame captured from a common webcam or PTZ camera.

Two main tasks are considered in this paper: (i) eye region detection and (ii) eye state recognition. The most fundamental task is to locate the eye regions when a human face is determined. Therefore, the eye state recognition can indicate some evidences of the subject (human). Let us consider an example in car driving. Camera systems have been installed in cars to monitor drivers' status in order to provide enough assistances if necessary. When the system finds that the driver's eye status keeps closed in a predefined time period, it can strongly imply that the driver is in drowsing. Similar functions can be also applied to our classrooms. The eye state recognition can automatically count the drowsing students to understand the attraction in course contents. This paper proposes a real-

time framework to efficiently locate the eye regions in a face image and determine the corresponding eye states.

Figure 1 draws the brief flowchart of our work in this paper. When a video frame is captured, a face detection method is first employed to demonstrate the face regions in image. Here we only consider a single frame of video, and the advanced applications should concatenate the multiple decisions of frames from a video sequences. We do not limit the number of faces appeared in image, but only request the resolution of eye regions. That is to say, the system has to guarantee the eye regions large enough (at least 50 pixels in width in our experiments) such that more information included for well recognition. Given a face region, all of pixels are converted to gray levels and then converted to binary. The binarization process for a face region can enhance the contrast between the pupil and its surrounding areas. Then, the horizontal and vertical projection that is simple but efficient is designed to locate the eye regions in the face region. Two kinds of visual features are extracted from eye regions, containing the intensity values of pixels and the local binary patterns (LBP) [3]. We tried to employ the two visual features individually with the support vector machine (SVM) [4] to recognize the eye states. Our experiments in this paper demonstrate the comparison of the efficiency of the two visual features used in the eye state recognition.
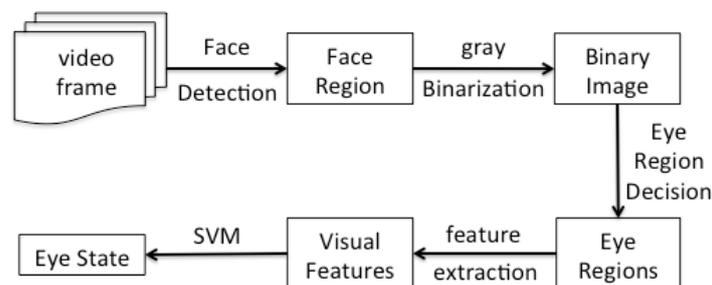

Figure 1: Flowchart

Our contribution in this paper mainly involves two parts. First, we propose a simple but efficient framework for eye state recognition. Our approach can achieve the real-time requirement such that less limitation is given in a real eye-related application. The proposed framework Moreover, each component in Figure 1 can be replaced to other more proper methods. For example, we can employ other visual features such as entropy or other informatics features to represent the eye regions for recognition. Second, this paper compares the efficiency of the two visual features (the intensity values of pixels and LBP) with the SVM in the eye region recognition. In our experiments, the visual representation using the intensity values of pixels can achieve a better performance than that using LBP in the eye state recognition.

The remainder of this paper is organized as the follows. Section 2 provides a brief review of the related works for the eye region detection and the eye state recognition. Section 3 presents our approach to detect eye regions in a face image, including face detection, image binarization, and eye region decision. Section 4 introduces the details of the eye state recognition used in our proposed framework. Section 5 first describes the experimental setup and data and then demonstrates the experimental results to show the performance of our proposed method. Section 6 makes a conclusion for this paper and provides possible extensions in the future.

2   Related Works

Many practical systems for eye tracking have been developed in market such as Tobii [1] and SMI [2]. How to well localize the eye regions in a face and then to identify the eye states are the two most fundamental issues in an eye tracking system.

Eyes locate at the fixed region on human face, where the eye regions are often darker than the surrounding skin mainly due to the low intensity of the pupil. Hence, an intuitive method to detect the eye regions is to employ the horizontal and vertical projection to locate the low intensity area in a binary face image [5]. Our proposed method to locate the eye regions basically follows the idea in [5] with some necessary modifications. R. Valenti and T. Gevers employed [6] the shape of the eye that is a significant feature in face to locate the eye regions. The pupil looks almost circle when an eye is open, so their approach sampled the pupil edges to identify the circle center that is also the eye center in face. Y. Wu et al. extracted the ellipse edges of eye as visual features for template matching to locate the eye regions [7]. Also, the OpenCV library included a template matching approach for eye region detection by using Haar-like features [8].

The eyes in general contain two states: open and closed, with the assumption of the eye regions available. J. Wu and M. M. Trivedi integrated the features in face to determine the eye states and to track the pupil movement [9]. H. Ta and Y.-J. Zhang recognized the eye states by looking for the iris of eye: iris found indicates the eye open [10]. Considering the variation of eye heights in open and closed, M. Dehnavi et al. performed the vertical project in the binary eye image to determine the eye open and closed with maximum and minimum thresholds, respectively [11]. C. Xu et al. adopted the LBP approach [12] for representing eye images and then construct an adaboosting classifier for eye state recognition [13]. Similarly, Y.-S. Wu et al. also employed the LBP approach but the SVM classification was adopted [14]. These two works presented the good performance of using the LBP feature for eye image representation.

3   Detection for Eye Regions

The first main issue in this paper is to detect the eye regions given a face image. This component shown in Figure 1 can be divided into three steps: face detection, conversion of the binary image, and eye region decision. This section describes the details of the three steps in the follows.

3.1   Face Detection

Face detection is a key technology to locate or identify human in image. Many researchers have paid attention to treating the face detection in either image or video under different conditions of the real world [15][16].

This work adopted the adaboosting approach [17], which was first proposed by Viola and Jones in 1999, to automatically detect face regions in a video frame. This approach collects a large number of Haar-like features that are fast computed by the integral image. Then, an adaboosting approach [18] is employed to select most significant Haar-like features of face regions. In general, many Haar-like features may be contained to construct an efficient classifier in the adaboosting approach, but that should also need more computational time for face detection. In order to achieve the real-time requirement, Viola and Jones designed a cascading structure, the basic concept is shown in Figure 2, to fast filter most of trivial non-face regions out using fewer Haar-like features in the several starting stages and then accurately to detect face regions using more features in the ending stages.
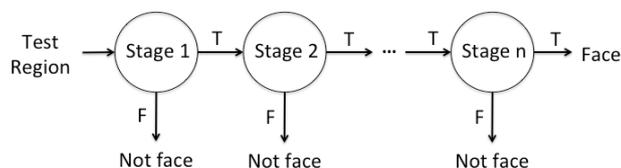

Figure 2: The cascade structure

The adaboosting approach of face detection was implemented in the openCV library and was widely used in many applications. Although this approach may be failed if the orientation of face regions is skew, it is still appropriated for our framework due to skewed human faces can be simply ignored in eye-related applications such as the eye state monitoring for car drivers.

3.2   Binarization

In face, the eye region contains a strong characteristic that there is high contrast between the eyes and the surrounding areas. Hence, the gray channel is more appropriate that color channels in representing eye regions. The other advantage of using the gray channel is to reduce the computational complexity in process.

This step of binarization contains two parts. The first is to convert the color to the gray image by using the simplest averaged method: gray = (R+G+B)/3, where R, G, and B are the pixel values in the red, green, and blue channels, respectively. Next, the gray image of face is thresholded to a binary image by using the Otsu method [19]. Figure 3 demonstrates an example from a color face image to a binary image.
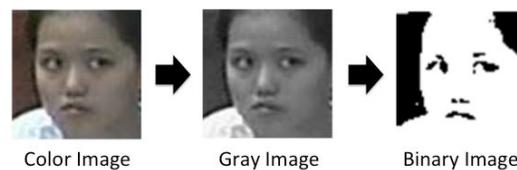

Color Image   Gray Image   Binary Image

Figure 3: A color face image to a binary image

3.3   Eye Region Decision

Many state-of-the-art technologies have been developed to eye region detection, stated as the above in Section 2. Since our work aims to design an eye state recognition for real-time applications, it is necessary to look for an efficient but fast approach to treat the eye region decision.

When a face has been located without skew, we can observe that the eye area is almost fixed on the upper-to-middle of the face. Hence, researchers defined a rough area in face to reduce the search space for the eye regions [20]. Let the width and the height of the face image be *w* and *h*, respectively. We averaged the position of eye in face in the training data, which is described in Section 5, to define that the left and right eyes individually locate at a rectangle region ((0.2, 0,23), (0.42, 0.5)), and ((0.58, 0.23), (0.8, 0.5)), respectively, in a face, where each of coordinates contains the left-upper and right-down corners. The candidate of eye regions defined in this work is drawn in Figure 4.
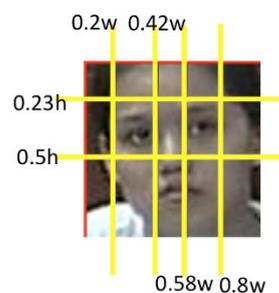

Figure 4: Candidate of two eye regions in a face

More advanced tasks are necessary to accurately extract the eye regions, such as ignoring the eyelashes and the shadow of eye. Due to our real-time requirement, we adopted the horizontal and vertical projection method [5] to fast but efficient capture the eye regions. Assume that $I(x, y)$ is the

pixel value at (x, y) in a candidate of eye region with size *m* by *n* pixels. Because the image is converted to binary, the pixel value is either 0 or 1. The horizontal (H_{porj}) and vertical (V_{proj}) projection on the binary image can be defined as the follow,

$$H_{proj}(x) = \sum_{x=1}^{n} I(x, y), \text{ where } x = 1 \text{ to } m$$

$$V_{proj}(y) = \sum_{x=1}^{m} I(x, y), \text{ where } y = 1 \text{ to } n.$$

An example of horizontal projection to align the horizontal center of an eye region is drawn in Figure 5. The distribution of the horizontal projection in Figure 5(c) contains two obvious peaks, where the upper peak corresponds to the eyelashes and the lower one to the eye region. Thus, we take the lower peak to align the horizontal center of the eye region. Moreover, the vertical projection to align the vertical center of the same eye region is drawn in Figure 6. Both in the closed and open eyes in this example, the peak of the vertical projection in Figure 6(c) can indicate the center of the eye region. Hence, the intersection of the horizontal and the vertical projection can exactly indicate the eye center.
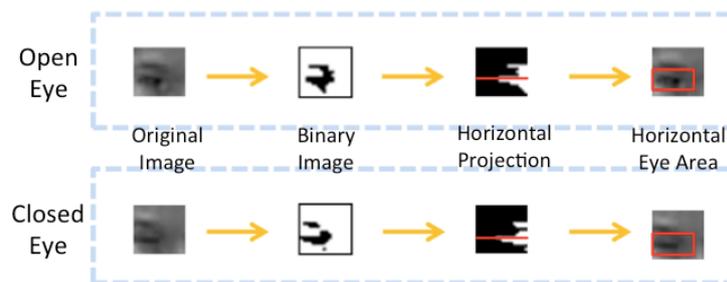

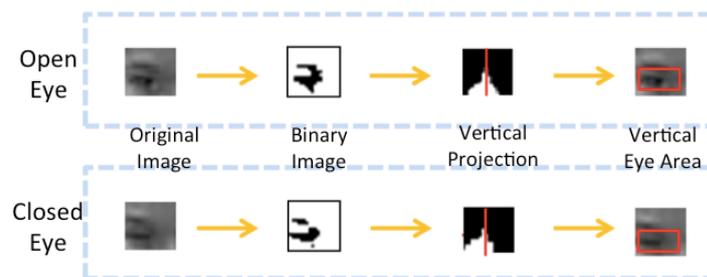**Figure 5: Example of the horizontal projection in an eye region**

Figure 6: Example of the vertical projection in an eye region

4   Eye State Recognition

The second main issue in this paper is to recognize the eye state given an eye region. To this end, a visual feature is extracted for representing the eye region and then a classification method is applied to recognize the eye state. Here only two eye states are considered: closed and open. In this paper, we adopted two kinds of visual features: pixel values and the LBP [3] in an eye region, and we employed the SVM approach [4] to the eye state classifier. In order to have the same length of feature vectors for the SVM classifier, we resize all eye regions (both training and test) to the equal size.

In our experiments described in the next section, we compare the performances of two visual features in the eye state recognition. The first one is to simply collect all values of row-major pixels in the eye region as a feature vector. In this case, all raw information in pixels is preserved without any encoding.

The second visual feature used in our experiments is the LBP method. The details of the LBP method are described in [12], and we only briefly introduce the concept in this section using the example in Figure 7. Assume that Figure 7(a) is a 3x3 cell in an image. Except the center value, all pixels are assigned a binary value shown in Figure 7(b): set 1 if the pixel value is greater than the center value, and set 0 otherwise. That is to say, the corresponding binary cell in Figure 7(b) remarks an intensity change of the cell associated with the center pixel. Moreover, a fixed weighted mask in Figure 7(c) encodes the binary cell to a LBP value in Figure 7(d). An image can then be converted to a series of LBP values by sliding 3x3 cells.

53	31	22
57	41	38
21	66	43

(a)

1	0	0
1	■	0
0	1	1

(b)

2^0	2^1	2^2
2^7	■	2^3
2^6	2^5	2^4

(c)
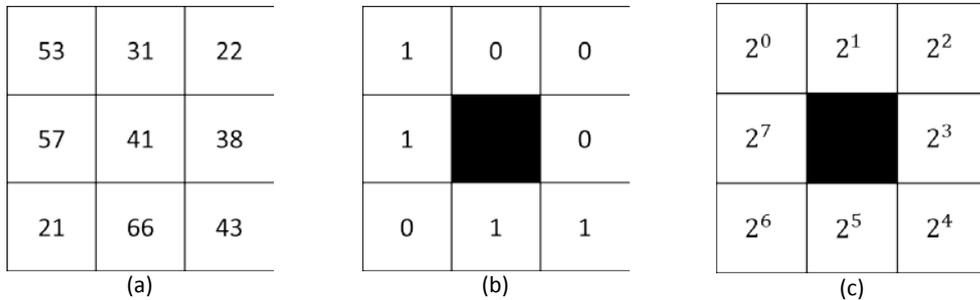
The LPB value of cell (a) is: $2^0+2^4+2^5+2^7=1+16+32+128=177$

(d)

Figure 7: LBP (a) a 3x3 cell, (b) the corresponding binary, (c) the fixed weights, and (d) the LBP value of (a)

Collecting visual features of the training eye regions, an SVM classifier is learnt to recognize test eye regions. SVM is a supervised learning model and is well-known to achieve a good performance in classification. In implementation, we adopted LIBSVM library [21] with a radial basis kernel. LIBSVM library was developed by Machine Learning and Data Mining Group, National Taiwan University, and can support the main functions of SVM classifiers. The SVM approach designs a linear classifier to best separate positive and negative data by the support vectors. However, a linear method is not able to well classify data in a complex problem, hence we adopt a radial basis kernel to project data into a very-high dimensional space such that the classification problem is separatable.

5   Experiments

5.1   Setup and Data Set

Figure 8 demonstrates the practical environments in our experiments that contain short and long distances to capture human eyes. In the both two cases, each of face images has to be larger than 60x60 pixels in order to catch enough information in eyes. All experiments are performed at a desktop that is with Intel(R) Core(TM) i7-3770: CPU @ 3.4GHz, RAM 8.0G, and windows 7.

Figure 8: Practical environments in our experiments. Left: shot distance; right: long distance.

Two data sets are collected in our experiments. The first data set is called D1 that asked for 17 people without glasses looking at the screen in the short distance in the left of Figure 8. There are six types of gazes in D1: front, up, down, right, left, and eye closed. In total, D1 contains 2260 eye images where 1500 for open and the other 760 for closed. The second data set is called D2 that was captured in classroom to record students' eye states in the right of Figure 8. This data set covers 39 people, 28 with glasses and the other 11 without glasses. Due to a long distance, the eye regions in D2 are low resolution. Table 1 summarizes the information of D1 and D2.

Table 1: Summary of our data sets D1 and D2

Data Set: D1	Open Eye					Closed Eye
Gaze	front	up	down	left	right	
# for training	100	100	100	100	100	500
# for test	200	200	200	200	200	260
total	300	300	300	300	300	760
Data Set: D2	Open Eye					Closed Eye
# for training	250					250
# for test	750					500
total	1000					750

5.2   Results

Our experiments mainly aim to evaluate the performance of our framework that applies the SVM approach as a classifier with two kinds of visual features: pixel values and the LBP. We adopted the LIBSVM library [21] with a radial basis kernel that is developed by Machine Learning and Data Mining Group, National Taiwan University. The number of training and test sets are listed in Table 1.

In the data set D1 that considers a short distance to capture eye images, we took 1000 images of the open and closed eye for training, and the others 1260 images for test. The recognition results using the two visual features are drawn in Table 2, which presents the pixel values can obviously achieve a better performance than LBP. Moreover, the LBP with SVM is faster than the pixel values with SVM.

Table 2: Recognition rates of the two visual features in data set D1.

	Open Eye (%)	Closed Eye (%)	Avg. Exec. Time (sec/per frame)
Pixel Values + SVM	99.8	100	0.0610
LPB + SVM	94.5	100	0.0401

Table 3: Recognition rates of the two visual features in data set D2

	With glasses			Without glasses		
	Open Eye (%)	Closed Eye (%)	Avg. Exec. Time (sec/per frame)	Open Eye (%)	Closed Eye (%)	Avg. Exec. Time (sec/per frame)
Pixel Values + SVM	94.53	88.26	0.0636	96.6	95.4	0.0542
LPB + SVM	82.6	86.9	0.0404	84.4	88.8	0. 0404

In the data set D2 that treat a long distance to capture low quality eye images, we took 500 images of the open and closed eye for training, and the others 1250 images for test. The recognition results using the two visual features are drawn in Table 3, which consists of with and without glasses. The results in Table 2 and 3 clearly present the better recognition rates using the feature of pixel values in both short and long distances.

6   Conclusion and Future Work

This paper presents a framework based on computer vision methods to treat eye region detection and eye state recognition. In the eye region detection, a simple but effective method is proposed to fast locate the eye regions in a face image. Then we employ two kinds of visual features, pixel values and LBP values, with SVM approach to classify the eye states. We also perform several experiments to demonstrate the performance of our proposed framework using these two visual features. The potential extension of this work can employ other efficient features including color, texture, and shape to improve the representation for eye regions. Another important task is to apply our proposed method to a practical system such as monitoring the driver status for car driving and recording the student status in classroom.

REFERENCES

[1].   Tobii, http://www.tobii.com/

[2].   SMI, http://www.smivision.com/en.html

[3].   Ojala, T., Pietikäinen, M., and Harwood, D., A comparative study of texture measures with classification based on feature distributions. Pattern Recognition, 1996. 19(3): p. 51-59.

[4].   Cortes, C. and Vapnik, V., Support-vector networks. Machine Learning, 1995. 20(3): p. 273-297.

[5].   Zhou Z.-H. and Geng X., Projection functions for eye detection. Pattern recognition, 2004. 37(5): p. 1049-1056.

[6].   Valenti R. and Gevers T., Accurate eye center location and tracking using isophote curvature. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2008. p. 1-8.

[7].   Wu Y., Liu H., and Zha H., A new method of detecting human eyelids based on deformable templates. Proceedings of IEEE Conference on Systems, Man and Cybernetics, 2004. p. 604-609.

[8].   OpenCV, http://opencv.org/

[9].   Wu J. and Trivedi M. M., An eye localization, tracking and blink pattern recognition system: algorithm and evaluation. ACM Transactions on Multimedia Computing, Communications, and Applications, 2010. 6(2): p. 599-601.

[10].   Tan H. and Zhang Y.-J., Detecting eye blink states by tracking iris and eyelids. Pattern Recognition Letters, 2006. 27(6): p. 667-675.

[11].   Dehnavi M., Attarzadeh N., and Eshghi M., Real time eye state recognition. Proceedings of the 19th Iranian Conference on Electrical Engineering, 2011. p. 1-4.

[12].   Ojala T., Pietikainen M., and Maenpaa T., Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002. 2497): p. 971-987.

[13].   Xu C., Zheng Y., and Wang Z., Efficient eye states detection in real-time for drowsy driving monitoring system. Proceedings of International Conference on Information and Automation, 2008. p. 170-174.

[14]. Wu Y.-S., Lee T.-W., Wu Q.-Z., and Liu H.-S., An eye state recognition method for drowsiness detection. Proceedings of IEEE Vehicular Technology Conference, 2010. p. 1-5.

[15]. Hsu R.-L., Abdel-Mottaleb M., and Jain A. K., Face detection in color images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002. 24(5): p. 696-706.

[16]. Liao S., Jain A. K., Li S. Z., Partial face recognition: alignment-free approach. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013. 35(5): p. 1193-1205.

[17]. Viola P. and Jones M. J., Robust real-time face detection. International journal of computer vision, 2004. 57(2): p. 137-154.

[18]. Freund, Y., Schapire, R., and Abe, N., A short introduction to boosting. Journal-Japanese Society For Artificial Intelligence, 1999. 14: p. 771-780.

[19]. Otsu N., A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man and Cybernetics, 1979. 9(1): p. 62-66.

[20]. Królak A. and Strumiłło P., Eye-blink controlled human-computer interface for the disabled. Advances in Intelligent and Soft Computing, 2009. 60: p. 123-133.

[21]. [LIBSVM] LIBSVM: http://www.csie.ntu.edu.tw/~cjlin/.

Powdery Mildew Disease Identification in Karpoori Variety of Betel Vine Plants using Histogram Based Techniques

[1]J.Vijayakumar and [2]Dr. S.Arumugam
[1]ASP/ECE, Nandha College of Technology, Erode, Tamil nadu, India.
[2] Principal, Nandha Engineering College, Erode, Tamil nadu, India.
vijiece@yahoo.co.in; arumugamdote@yahoo.co.in

ABSTRACT

 The betel vine cultivation is very much affected by diseases and outcome of the farmer is big loss for betel vine cultivation. The aim of this paper is to detection of Powdery mildew disease in the betel vine plants using histogram based digital image processing techniques. The digital images of the uninfected or normal betel vine leaves and the digital images of the infected in powdery mildew diseased betel vine leaves at different stages are collected from different betel vine plants using a high resolution digital camera and collected betel vine images are stored with JPEG format. The digital image analyses of the leaves are done using the image processing toolbox in MATLAB. The RGB color betel vine images were converted into gray scale image. Histogram were plotted and stored as a database for uninfected and powdery mildew disease infected in first day to final day for all the Karpoori variety of betel vine leaves. Finally this investigation helps to recognize the powdery mildew disease can be identified before it spreads to entire crop.

Keywords: Piperaceae; Piper betel; Powdery Mildew disease; Oidium Piperis

1   Introduction

The betel vine leaves were popularly known as Vettilai in Tamil and also commonly known as Paan in Hindi. Biological name of betel vine leaves were known as Piper betel. It belongs to the family of Piperaceae. The Vitamins B and C is highly available in the betel vine leaves and they were mainly used in a tonic to the brain, liver and heart for human [95]. The fresh juice of betel vine leaves are used to many ayurvedic preparations. The betel vine plants were cultivated throughout India except the dry northwestern parts. The six betel vine leaves with a little bit of slaked lime are equal to 300 ml of cow milk particularly for the vitamin and mineral nutrition. The group of research work was going on in the field of betel vine disease analysis for various centers within the country under the name "All India Coordinated Research Project on Betel vine". 70 varieties of betel vine leaves are Cultivated in the world. Among these 70 varieties, 40 varieties of betel vine leaves are Cultivated in India. 30 varieties of betel vine leaves are Cultivated in West Bengal. Tamil Nadu, Uttar Pradesh, Bihar, Maharashtra, Karnataka, West Bengal, Andhra Pradesh and Kerala states are widely cultivated in the betel vine. In Tamilnadu, based on the color, size and taste, there are many varieties of betel vine leaves available and some of the most popular varieties are vellaikodi, Karpoori, pachaikodi and Sirugamani. Karpoori variety of betel vine leaves was considered in this research work paper. During the cultivation of betel vine, diseases were one of the most important causes that reduce quantity of the betel vine leaves. The most important betel vine plants diseases were powdery mildew disease, leaf rot disease, foot rot disease and leaf spot disease.

Powdery mildew disease is caused by Oidium piperis. The disease appears on the undersurface of the leaves as white to brown powdery patches. These patches gradually increase in size and often coalesce with each other. They vary in size from a few to 40mm in diameter and are covered by dusty growth which is fairly thick in cases of severe attack [95][96]. Areas on the upper surface corresponding to patches on the under surface appear yellowish, raised and irregular in outline. Young leaves when infected fail to grow and become deformed the surface being cracked and the margin turned inwards. The disease has been reported to be in the leaves only and it has been found to
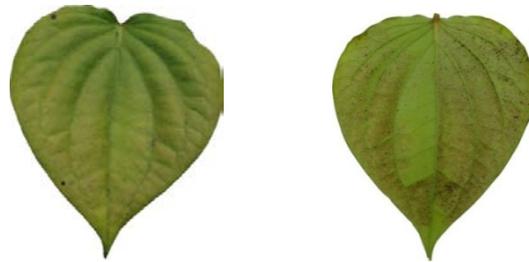
Figure 1. Powdery mildew disease affected betel vine leaves

disappear during the hot season. Figure 1 shows the images of front and back view of powdery mildew infected betel vine leaves. In this research paper Powdery mildew disease were considered for Karpoori variety of betel vine plants.

2 Materials and Methods

For the Histogram based analysis, the front and back view normal or healthy betel vine leaves and powdery mildew disease infected in early stage to final stage betel vine leaves were individually collected at different plants using a high-resolution digital camera for all Karpoori betel vine plants from erode, karur and trichy district of Tamil nadu, India. The collected betel vine leaves back grounds were eliminated using photo shop and these digital images were stored in the system [103]. These stored betel vine images were resized. Digital imaging techniques were divided into two phases. Normal or uninfected betel vine leaves phase and Powdery mildew disease infected in early stage to final stage of betel vine leaves phase. The uninfected betel vine leaves phase consists of without any disease infected in the betel vine leaves. The same variety of front and back view betel vine leaves were collected at different betel vine plants and place. The RGB color betel vine images were converted into gray scale image. Histogram were plotted and stored as a database for all Karpoori variety of betel vine leaves.

Infected betel vine leaves phase consists of Powdery mildew disease infected in early stage to final stage of betel vine leaves. The same varieties of betel vine sample leaves were selected from uninfected betel vine plant, which is nearest to the betel vine plant infected by Powdery mildew disease. The serial numbers were given to all selected betel vine sample leaves. The front and back view betel vine sample leaves was collected serial number wise at different plants and place. The RGB color betel vine images were converted into gray scale image. Histogram was plotted for all Karpoori variety of betel vine leaves and plotted histogram were compared with the stored data base values. If the calculated Histogram plot and stored Histogram plot were same range for all gray scale values, the selected betel vine leaves were included in samples otherwise samples were removed to the selected list.

The accepted same variety of betel vine sample leaves were collected serial number wise for next two three days. The RGB color betel vine images were converted into gray scale image. Histogram was plotted for all Karpoori variety of betel vine leaves and plotted histogram were compared with the stored data base values. If any difference were identified between calculated and stored database values on any particular day for the particular betel vine leaf, that particular day were counted as Powdery mildew disease infected in first day for the particular betel vine sample leaf. These Powdery mildew disease infected betel vine leaves were collected serial number wise for infected in first day to final day. The RGB color betel vine images were converted into gray scale image. Histogram was plotted for all Karpoori variety of betel vine leaves and plotted histogram were stored data base values.

3   Result

The histogram for front and back view normal betel vine leaves was shown in figure 2.The histogram for front and back view powdery mildew disease infected in first day betel vine leaves were shown in figure 3.The histogram for front and back view powdery mildew disease infected in second day betel vine leaves were shown in figure 4.The histogram for front and back view powdery mildew disease infected in third day betel vine leaves were shown in figure 5.The histogram for front and back view powdery mildew disease infected in fourth day betel vine leaves were shown in figure 6.The histogram for front and back view powdery mildew disease infected in fifth day betel vine leaves were shown in figure 7.

The gray scale value of the histogram for front view of uninfected betel vine leaf was between 100 and 250. However the initial gray scale value was near to 100. Final gray scale value was near to 200. Maximum frequency of occurrence of the gray scale value was between 150 and 200.The gray scale value of the histogram for back view of uninfected betel vine leaf was between 150 and 250. However the initial gray scale value was near to 150. Final gray scale value was near to 250. The maximum frequency of occurrence of the gray scale value was near to 200.The gray scale value of the histogram for front view of betel vine leaf with powdery mildew disease at first day of infection was between 50 and 200. However the initial gray scale value was near to 100 and final gray scale value was near to 150.
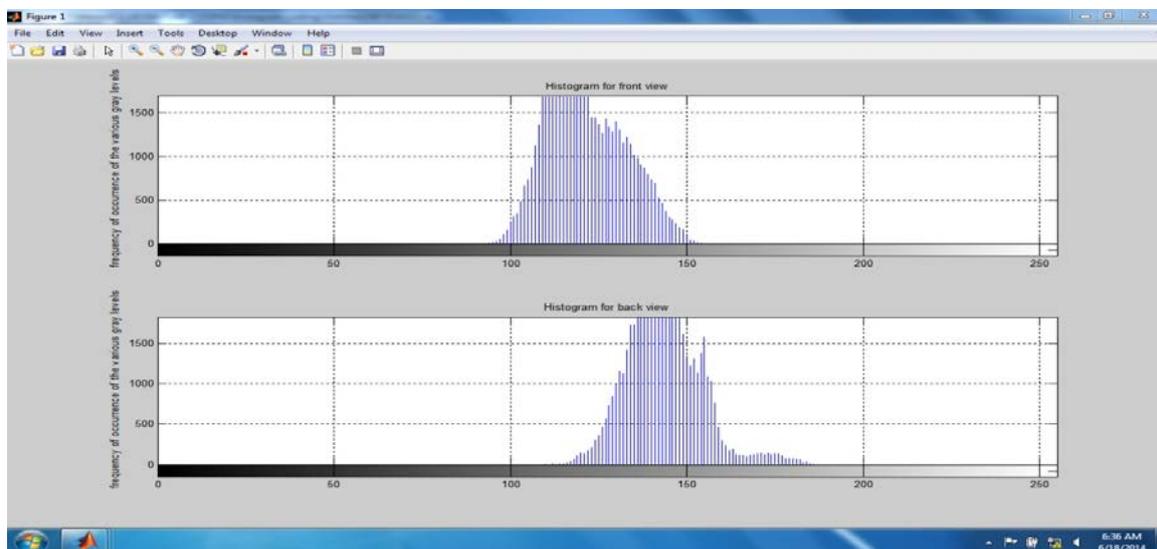

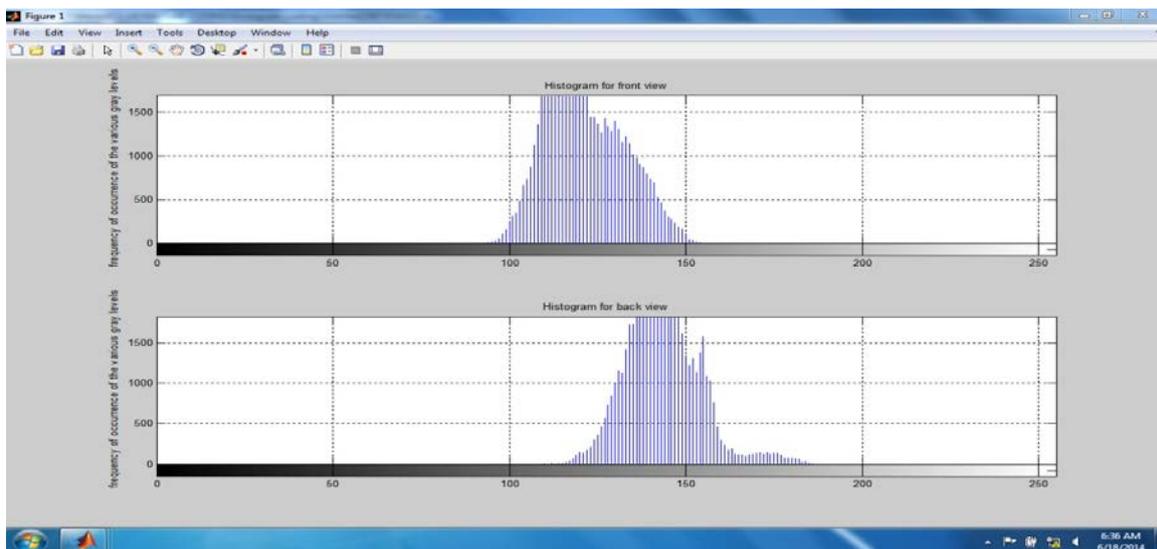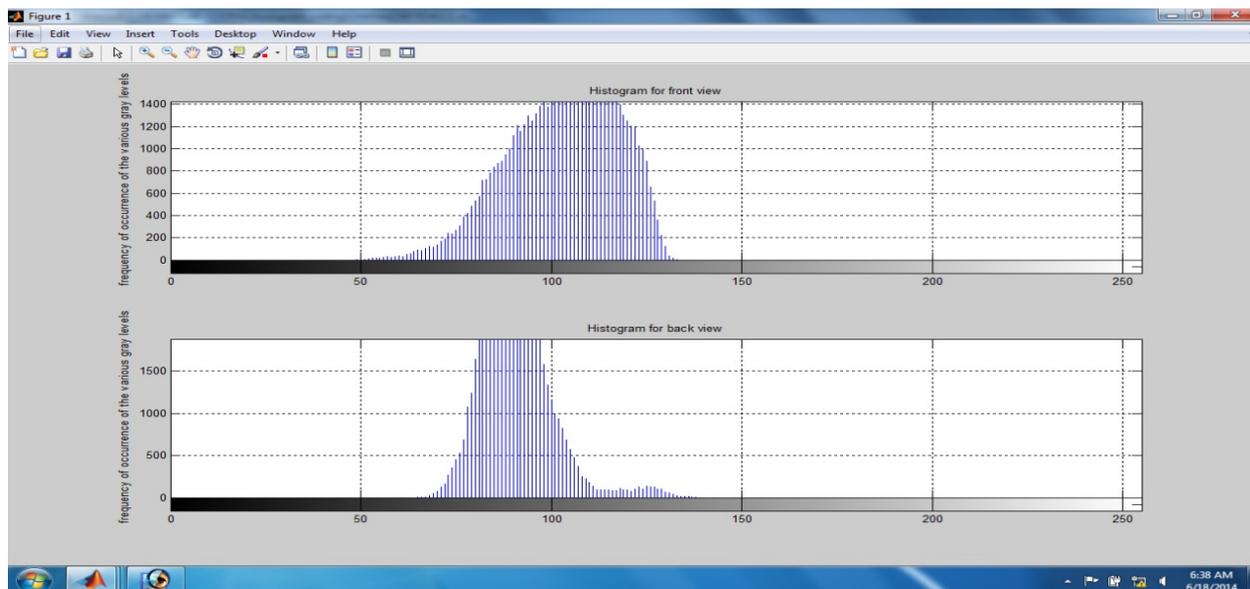Figure 2. Histogram for front and back view normal betel vine leaves

Figure 3. Histogram for front and back view powdery mildew disease infected in first day betel vine leaves

Maximum frequency of occurrence of the gray scale value was between 100 and 150.The gray scale value of the histogram for back view of betel vine leaf with powdery mildew disease at first day of infection was between 100 and 200. However the initial gray scale value was near to 100 and final gray scale value was near to 150. Maximum frequency of occurrence of the gray scale value was near to 150.The gray scale value of the histogram for front view of betel vine leaf with powdery mildew disease at second day of infection was between 0 and 150. However the initial gray scale value was near to 50 and final gray scale value was near to 150. Maximum frequency of occurrence of the gray scale value was near to 100.The gray scale value of the histogram for
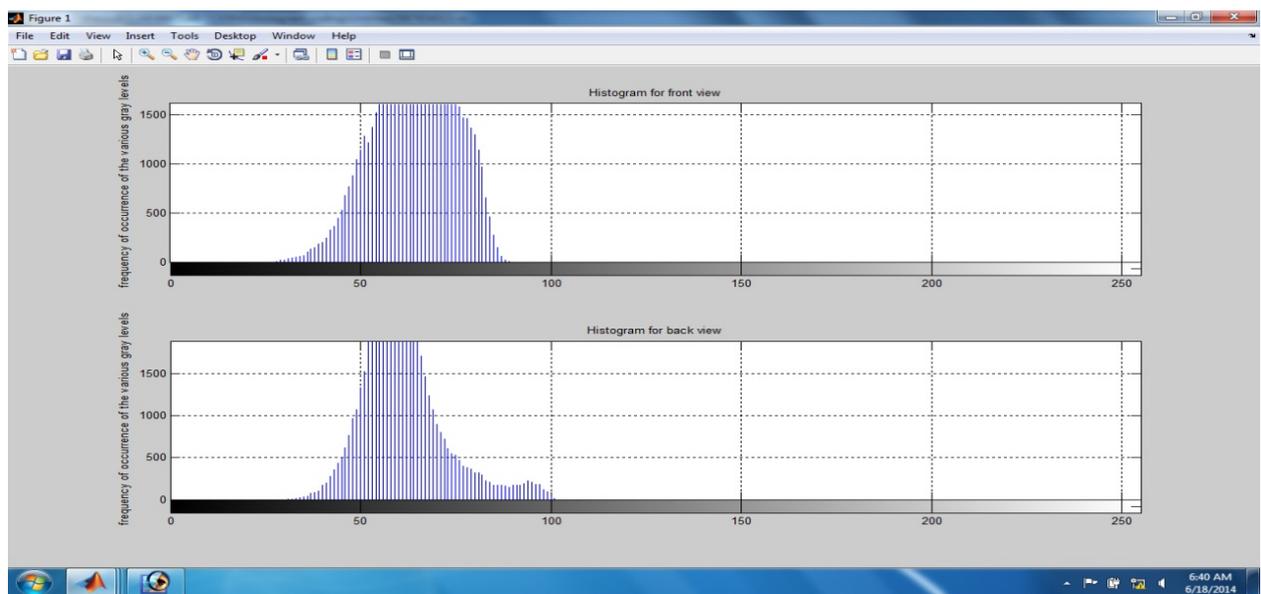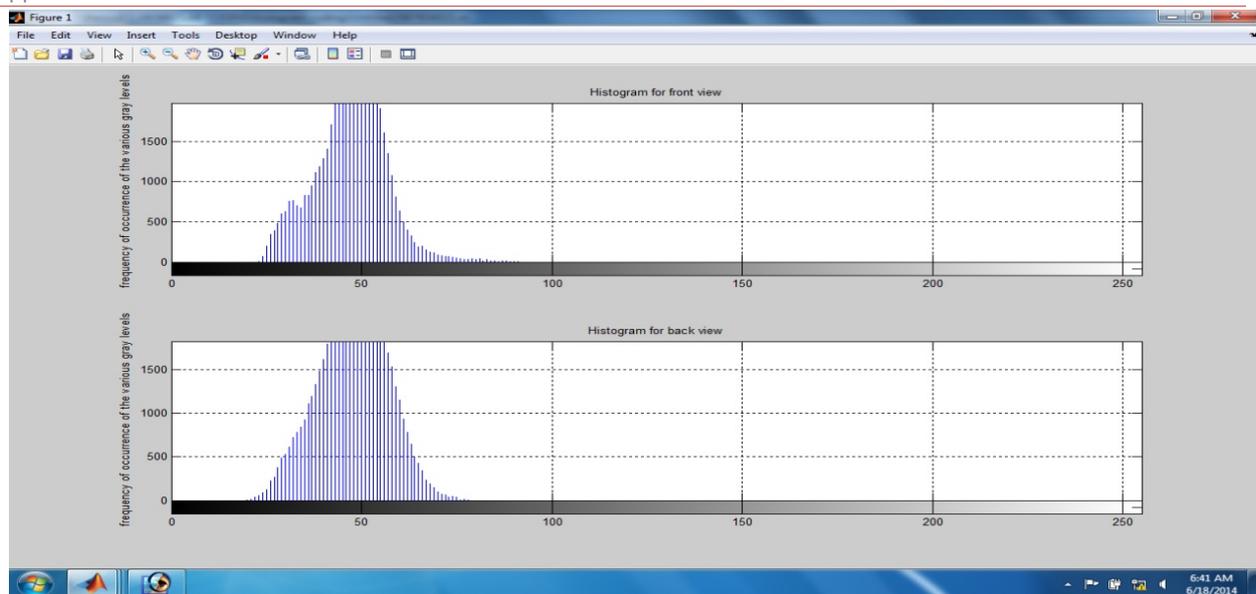

Figure 4. Histogram for front and back view powdery mildew disease infected in second day betel vine leaves
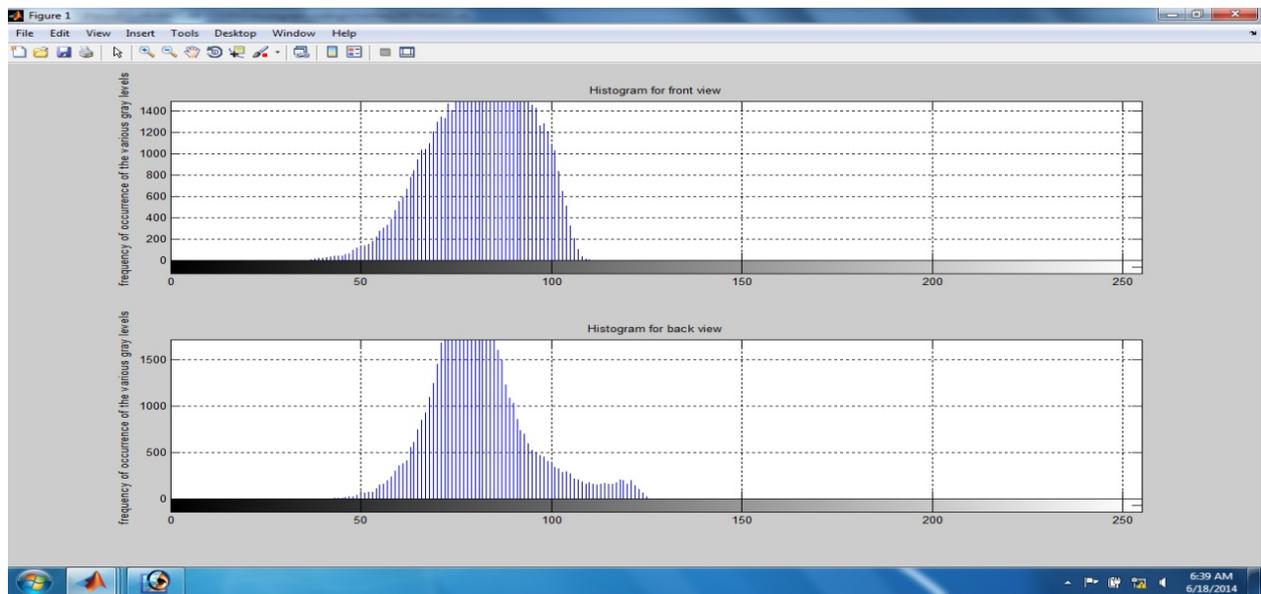
Figure 5. Histogram for front and back view powdery mildew disease infected in third day betel vine leaves

back view of betel vine leaf with powdery mildew disease at second day of infection was between 50 and 150. However the initial gray scale value was near to 50 and final gray scale value was near to 150. Maximum frequency of occurrence of the gray scale value was near to 100. The gray scale value of the histogram for front view of betel vine leaf with powdery mildew disease at third day of infection was between 0 and 150. However the initial gray scale value was near to 50 and final gray scale value was near to 100. Maximum frequency of occurrence of the gray scale value was between 50 and 100.The gray scale value of the histogram for back view of betel vine leaf with powdery mildew disease at third day of infection was between 50 and 150. However the initial gray scale value was near to 50 and final gray scale value was between 100 and 150.

Figure 6. Histogram for front and back view foot rot disease infected in fourth day betel vine leaves

Figure 7. Histogram for front and back view powdery mildew disease infected in fifth day betel vine leaves

Maximum frequency of occurrence of the gray scale value was between 50 and 100.The gray scale value of the histogram for front view of betel vine leaf with powdery mildew disease at fourth day of infection was between 50 and 100. However the initial gray scale value was between 0 and 50 and final gray scale value was near to 100. Maximum frequency of occurrence of the gray scale value was between 50 and 100.The gray scale value of the histogram for back view of betel vine leaf with powdery mildew disease at fourth day of infection was between 50 and 100. However the initial gray scale value was between 0 and 50 and final gray scale value was 100. Maximum frequency of occurrence of the gray scale value was near to 50.The gray scale value of the histogram for front view of betel vine leaf with powdery mildew disease at fifth day of infection was between 0 and 100. However the initial gray scale value was between 0 and 50 and final gray scale value was near to 100. Maximum frequency of occurrence of the gray scale value was near to 50.The gray scale value of the histogram for back view of betel vine leaf with powdery mildew disease at fifth day of infection was between 0 and 100. However the initial gray scale value was between 0 and 50 and final gray scale value was near to 100. Maximum frequency of occurrence of the gray scale value was near to 50.

4   Conclusion

The above research techniques express that the karpoori variety of betel vine plants Oidium piperis fungus can be recognized in starting stage of betel vine plantation and saved before the Oidium piperis fungus starts to reach complete karpoori variety of betel vine crop. The method of detecting the Oidium piperis fungus cost is effective. This technique can also be extended to detect fungus or diseases of all kind plants to recognize starting stage preventive action.

REFERENCES

[1].   Alham F. Aji, Qorib Munajat, Ardhi P. Pratama, Hafizh Kalamullah, Aprinaldi, Jodi Setiyawan & Aniati M. Arymurthy 2013, 'Detection of Palm Oil Leaf Disease with Image Processing and Neural etwork Classification on Mobile Device', International Journal of Computer Theory and Engineering, Vol.5, no.3, pp. 528-532.

[2].   Ajay A. Gurjar, Viraj A. Gulhane 2011, 'Disease Detection on Cotton Leaves by Eigenfeature Regularization and Extraction Technique' International Journal of Electronics, Communication & Soft Computing Science and Engineering, Vol.1, no.1, pp. 1-4.

[3].   Arunkumar Beyyala1 & Sai Priya Beyyala 2012, 'Application for Diagnosis of Diseases In Crops Using Image Processing', International Journal of Life Sciences Biotechnology and Pharma Research, Vol.1, no.1, pp. 172 -175.

[4].   S. Ananthi & S. Vishnu Varthini 2012, 'Detection and Classification of Plant Leaf Diseases', International Journal of Research in Engineering & Applied Sciences, Vol.3, no.1, pp. 763 -773.

[5].   S. Arivazhagan, R. Newlin Shebiah, S. Ananthi & S. Vishnu Varthini 2013, 'Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features', Agricultural Engineering International: CIGR Journal, Vol.15, no.1, pp. 211 -217.

[6].   Anand.H.Kulkarni & Ashwin Patil R. K 2012, 'Applying image processing technique to detect plant diseases', International Journal of Modern Engineering Research, Vol.2, no.5, pp. 3661 -3664.

[7].   Arti N. Rathod, Bhavesh Tanawal & Vatsal Shah 2013, Image Processing Techniques for Detection of Leaf Disease', International Journal of Advanced Research in Computer Science and Software Engineering, Vol.3, no.11, pp. 397 -399.

[8].   Amina Khatra 2013, 'A Novel Machine Vision System for Identification of Yellow Rust in Wheat Plants', International Journal of Image Processing, Vol.7, no.5, pp.430-435.

[9].   Arun Kumar.T & Karun Verma 2010, 'A Theory Based on Conversion of RGB image to Gray image', International Journal of Computer Applications, Vol.7, no.2, pp.7-9.

[10].   Alberto Tellaeche, Xavier P. Burgos-Artizzu, Gonzalo Pajares & Angela Ribeiro 2008, 'Avision-based method for weeds identification through the Bayesian decision theory', The journal of the Pattern Recognition society, Vol.4, no.1, pp.521-530.

[11].   Amir Alipasandi, Hosein Ghaffari & Saman Zohrabi Alibeyglu 2013, 'Classification of three Varieties of Peach Fruit Using Artificial Neural Network Assisted with Image Processing Techniques', International Journal of Agronomy and Plant Production, Vol.4, no.9, pp.2179-2186.

[12].   Al-Hiary.H, S. Bani-Ahmad, M. Reyalat, M. Braik & Z. ALRahamneh 2011, 'Fast and Accurate Detection and Classification of Plant Diseases', International Journal of Computer Applications, Vol.17, no.1, pp. 31 - 38.

[13].   Ajit Danti, Manohar Madgi & Basavaraj S. Anami 2012, 'Mean and Range Color Features Based Identification of Common Indian Leafy Vegetables', International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol.5, no.3, pp. 151 - 160.

[14].   Anjali Verma, Nikhil Kumar & S. A. Ranade 2004, 'Genetic diversity amongst landraces of a dioecious vegetatively propagated plant, betelvine (Piper betle L.)', Indian Academy of Sciences, Vol.29, no.3, pp. 319 -328.

[15].   Balamurugan.P & R. Rajesh 2012, 'Neural Network Based System for the Classification of Leaf Rot Disease in Cocos Nucifera Tree Leaves', European Journal of Scientific Research, Vol.88, no.1, pp. 137 - 145.

[16].   Basavaraj S. Anami, Suvarna S. Nandyal & A. Govardhan 2010, 'A Combined Color, Texture and Edge Features Based Approach for Identification and Classification of Indian Medicinal Plants', International Journal of Computer Applications, Vol.6, no.12, pp.45-51.

[17].   Basvaraj .S. Anami, J.D. Pujari & Rajesh.Yakkundimath 2011, Identification and Classification of Normal and Affected Agriculture/horticulture Produce Based on Combined Color and Texture Feature Extraction, Vol.1, no.3, pp. 356-360.

[18].   Balaji Kaveti, Lisa Tan, Sarnnia, Tan Sin Kuan & Mirza Baig 2011, 'Antibacterial Activity Of Piper Betel Leaves', International Journal of Pharmacy Teaching & Practices, Vol.2, no.3, pp. 129 -132.

[19].   Bibekananda Mohanty, Partha Datta, B. Dasgupta1 & Dalim Kumar Sengupta 2011, 'Integrated Management of Foot and Leaf Rot of Betelvine', SAARC Journal of Agriculture, Vol.9, no.2, pp. 83 -91.

[20].   Bindu Tiger & Toran Verma 2013, 'Identification and Classification of Normal and Infected Apples using Neural Network', International Journal of Science and Research, Vol.2, no.6, pp. 160 -163.

[21].   Burks T.F, A. Shearer S.A, & Payne F.A 2010, 'Classification of Weed Species Using Color Texture Features And Discriminant Analysis', American Society of Agricultural Engineers', Vol.43, no.2, pp.441-448.

[22].   P. Chowdappa,D. Brayford,J. Smith & J. Flood 2003, 'Identification of Phytophthora species affecting plantation crops by RFLP of PCR-amplified internal transcribed spacer regions of ribosomal RNA', Current Science, Vol.85, no.1, pp. 34 -36.

[23].   Diao Zhihua, Wang Huan, Song Yinmao & Wang Yunpeng 2013, 'Image Segmentation Method for Cotton Mite Disease Based on Color Features and Area Thresholding', Journal of Theoretical and Applied Information Technology, Vol.48, no.1, pp. 527 -533.

[24].   B. Dasgupta, B.Mohanty, P. K. DUTTA & Satyabrata Maiti 2008, 'Phytophthora Diseases of Betelvine (Piper Betle L.) – A Menace To Betelvine Crop', SAARC Journal of Agriculture, Vol.6, no.1, pp. 1 -19.

[25].   Debabrata Samanta & Arya Ghosh 2012, 'Histogram Approach for Detection of Maize Leaf Damage', International Journal of Computer Science and Telecommunications, Vol. 3, no. 2, pp. 26-28.

[26].   Dinesh Chandra Khatua, Bholanath Mondal & Rana Bhattachayya 2013, 'A selective medium for Xanthomonas axonopodis pv. Betlicola, bacterial pathogen of betelvine', African Journal of Agricultural Research, Vol.8, no.49, pp. 6388 -6393.

[27].   Faithpraise Fina, Philip Birch, Rupert Young, J. Obu, Bassey Faithpraise & Chris Chatwin 2013, 'Automatic Plant Pest Detection and Recognition Using K-Means Clustering Algorithm and Correspondence Filters', International Journal of Advanced Biotechnology and Research, Vol.4, no.2, pp. 189 -199.

[28].   A. R. Fathilah 2011, 'Piper betle L. and Psidium uajava L. in oral health maintenance' Journal of Medicinal Plants Research, Vol.5, no.2, pp. 156 -163.

[29].   Faisal Ahmed, Hasanul Kabir, Shayla Bhuyan, Hossain Bari & Emam Hossain 2014, 'Automated Weed Classification with Local Pattern-Based Texture Descriptors', The International Arab Journal of Information Technology, Vol.11, no.1, pp.87-94.

[30].   Faisal Ahmed, A.S.M. Hossain Bari, Emam Hossain, Hawlader Abdullah Al-Mamun & Paul Kwan 2011, 'Performance Analysis of Support Vector Machine and Bayesian Classifier for Crop and Weed Classification from Digital Images', World Applied Sciences Journal, Vol.12, no.4, pp.432-440.

[31].   Gayatri Vijay Patil Pranali vivekanand More, Nandinee Dattatrey Pawar 2014, 'SVM approach for color image classification using shape and texture', International Journal of Emerging Trends in Engineering and Development, Vol.3, no.4, pp. 139 - 146.

[32]. J. L. Gonzalez-Perez, M. C. Espino-Gudino, J. Gudino-Bazaldua, J. L. Rojas-Renteria, V. Rodriguez-Hernandez & V.M. Castano 2013, 'Color image segmentation using perceptual spaces through applets for determining and preventing diseases in chili peppers', African Journal of Biotechnology, Vol.12, no.7, pp. 679 -688.

[33]. P. Guha 2006, 'Betel Leaf: The Neglected Green Gold of India', Journal of Human Ecology, Vol.19, no.2, pp. 87 -93.

[34]. B. K. Goswamil, K. A. Kader, M. L. Rahman, M. R. Islam & P. K. Malaker 2002, 'Development of Leaf Spot of Betel Vine Caused By Colletotrichum Capsici', Bangladesh Journal of Plant Pathology, Vol.18, no.1, pp. 39 -41.

[35]. Guillaume Cerutti,Laure Tougneb, Julien Mille, Antoine Vacavant & Didier Coquin 2013, 'Understanding Leaves in Natural Images -A Model-Based Approach for Tree Species Identification' Computer Vision and Image Understanding, Vol.117, no.10, pp. 1482 - 1501.

[36]. V. A. Gulhane & A. A. Gurjar 2011, 'Detection of Diseases on Cotton Leaves and Its Possible Diagnosis', International Journal of Image Processing, Vol.5, no.5, pp. 590-598.

[37]. M I. Huq 2011, 'Studies on the Epidemiology of Leaf Rot and Leaf Spot Diseases of Betel Vine (Piper Betle L.)', Bangladesh Journal of Scientific and Industrial Research, Vol.46, no.4, pp. 519 -522.

[38]. Hong Fang & Huijie Li 2014, 'Plant Leaves Recognition and Classification Model Based on Image Features and Neural Network', IJCSI International Journal of Computer Science, Vol.11, no.1, pp. 100 - 104.

[39]. Hrishikesh P. Kanjalkar & S.S.Lokhande 2014,'Feature Extraction of Leaf Diseases', International Journal of Advanced Research in Computer Engineering & Technology, Vol.3, no.1, pp. 153 -155.

[40]. Jagadeesh Devdas Pujari, Rajesh Yakkundimath & Abdulmunaf Syedhusain Byadgi 2013, 'Grading and Classification of Anthracnose Fungal Disease of Fruits based on Statistical Texture Features', International Journal of Advanced Science and Technology, Vol.52, no.1, pp. 121 -132.

[41]. Jayamala K. Patil & Raj Kumar 2011, 'Advances In Image Processing for Detection of Plant Diseases', Journal of Advanced Bioinformatics Applications and Research, Vol.2, no.2, pp.135-141.

[42]. Junlong Fang，Shuwen Wang & Changli Zhang 2005, 'Application of Genetic Algorithm（GA）Trained Artificial Neural Network to Identify Tomatoes with Physiological Diseases', Nature and Science, Vol.3, no.2, pp.1-7.

[43]. V. V. Kaleeswari & T.Sridhar 2013, 'A Study on Betel Vine Cultivation and Market Crisis in Karur District', Indian Journal of Applied Research Vol. 3, no. 10, pp. 1-3.

[44]. Kiran. R, Gavhale & Ujwalla Gawande 2014, 'An Overview of the Research on Plant Leaves Disease detection using Image Processing Techniques', IOSR Journal of Computer Engineering, Vol.16, no.1, pp. 10 - 16.

[45]. Lakshmi Arambewela,K,G.A. Kumarratunga & Kalyani Dias 2005, 'Studies of Piper Betle of Sri Lanka', Journal of the National Science Foundation of Sri Lanka, Vol.33, no.2, pp. 133 -139.

[46]. Manoj P Rai, Karadka Ramdas Thilakchand, Princy L Palatty, Prathima Rao, Suresh Rao, Harshith P Bhat & Manjeshwar Shrinath Baliga 2011, 'Piper Betel Linn (Betel Vine), the Maligned Southeast Asian Medicinal Plant Possesses Cancer Preventive Effects: Time to Reconsider the Wronged Opinion', Asian Pacific Journal of Cancer Prevention, Vol.12, no.1, pp. 2149 -2156.

[47].   Manoj Mukherjee, Titan Pal & Debabrata Samanta 2012, 'Damaged Paddy Leaf Detection Using Image Processing', Journal of Global Research in Computer Science, Vol.3, no.10, pp. 7-10.

[48].   Maliheh Shabanzade, Morteza Zahedi & Seyyed Amin Aghvami 2011, 'Combination of Local Descriptors And Global Features For Leaf Recognition', Signal & Image Processing : An International Journal, Vol.2, no.3, pp. 23 - 31.

[49].   Mokhled S. Al-Tarawneh 2013, 'An Empirical Investigation of Olive Leave Spot Disease Using Auto-Cropping Segmentation and Fuzzy C-Means Classification' World Applied Sciences Journal, Vol.23, no.9, pp. 1207 - 1211.

[50].   Manpreet Kaur, Jasdeep Kaur & Jappreet Kaur 2011, 'Survey of Contrast Enhancement Techniques based on Histogram Equalization', International Journal of Advanced Computer Science and Applications, Vol.2, no.7, pp.137-141.

[51].   Mohamad Faizal Ab Jabal, Suhardi Hamid, Salehuddin Shuib & Illiasaak Ahmad 2013, 'Leaf Features Extraction and Recognition Approaches to Classify Plant', Journal of Computer Science, Vol.9, no.10, pp.1295-1304.

[52].   Madhuri V & Amrutha Gayathri D 2014, 'Root Rot of Chilli Incited By Sclerotium Rolfsii Sacc. And Its Management- A Review', International Journal of Applied Biology and Pharmaceutical Technology, Vol.5, no.1, pp.197-204.

[53].   Manimekalai .K & Vijaya.MS 2014, 'Support Vector Machine Based Tool For Plant Species Taxonomic Classification', Journal of Asian scientific research, Vol.4, no.4, pp.159-173.

[54].   T. Nalina & Z.H.A. Rahim 2007, 'The Crude Aqueous Extract of Piper betle L. and its Antibacterial Effect Towards Streptococcus mutans', American Journal of Biotechnology and Biochemistry, Vol.3, no.5, pp. 10 -15.

[55].   Narendra V G & Hareesh K S 2010, 'Prospects of Computer Vision Automated Grading and Sorting Systems in Agricultural and Food Products for Quality Evaluation', International Journal of Computer Applications, Vol.1, no.4, pp.1-9.

[56].   Narendra V G & Hareesh K S 2010, 'Quality Inspection and Grading of Agricultural and Food Products by Computer Vision- A Review', International Journal of Computer Applications, Vol.2, no.1, pp.43-65.

[57].   Nitin S. Tijare & Sagar S. Badnerkar 2014, 'Image Recognition Based Crop Disease Identification System: A Survey', International Journal of Computer Science and Mobile Computing, Vol.3, no.4, pp.868-873.

[58].   Namita Mittal, Basant Agarwal, Ajay Gupta & Hemant Madhur 2014, 'Icon Based Information Retrieval and Disease Identification in Agriculture', International Journal of Advanced Studies in Computer Science & Engineering, Vol.3, no.3, pp. 26 - 31.

[59].   Pranjali Vinayak Keskar, Shubhangi Nimba Masare, Manjusha Suresh Kadam & Seema U.Deoghare 2013, 'Leaf Disease Detection and Diagnosis', International Journal of Emerging Trends in Electrical and Electronics, Vol.2, no.2, pp. 28 - 31.

[60].   S. R. Pokharkar & V. R.Thool 2012, 'Early Pest Identification in Greenhouse Crops using Image Processing Techniques', International Journal of Computer Science and Network, Vol.1, no.3, pp. 1-6.

[61].   S. Phadikar, J. Sil & A. K. Das 2012, 'Classification of Rice Leaf Diseases Based on Morphological Changes, International Journal of Information and Electronics Engineering, Vol.2, no.3, pp. 460-463

[62].   K. Padmavathi 2012, 'Investigation and monitoring for leaves disease detection and evaluation using image processing' International Research Journal of Engineering Science, Technology and Innovation, Vol.1, no.3, pp. 66-70.

[63].   Piyush Chaudhary, Anand K. Chaudhari,A. N. Cheeran& Sharda Godara 2012, 'Color Transform Based Approach for Disease Spot Detection on Plant Leaf', International Journal of Computer Science and Telecommunications, Vol.1, no.6, pp. 65 -70.

[64].   Pradnya Ravindra Narvekar, Mahesh Manik Kumbhar & S. N. Patil 2013, 'Grape Leaf Diseases Detection & Analysis using SGDM Matrix Method', International Journal of Innovative Research in Computer and Communication Engineering. Vol.2, no.3, pp. 3365 - 3372.

[65].   Qinghai He, Benxue Ma, Duanyang Qu, Qiang Zhang,Xinmin Hou & Jing Zhao 2013, 'Cotton Pests and Diseases Detection based on Image Processing',Telkomnika Indonesian Journal of Electrical Engineering, Vol.11, no.6, pp. 3445 - 3450.

[66].   Rafiqul Haider.M,Abul Khair,M.Mahfuzur Rahman & M.Khairul Alam 2013, 'Indigenous Management Practices of Betel-Leaf (Piper Betle L.) Cultivation by the Khasia community in Bangladesh' Indian journal of Traditional knowledge, Vol.12, no.2, pp. 231 -239.

[67].   Rupesh G. Mundada & V. V. Gohokar 2013, Detection and Classification of Pests in Greenhouse Using Image Processing', IOSR Journal of Electronics and Communication Engineering, Vol.5, no.6, pp. 57-63.

[68].   Rothe P.R & Kshirsagar R.V 2014, 'SVM-based Classifier System for Recognition of Cotton Leaf Diseases', International Journal of Emerging Technologies in Computational and Applied Sciences, Vol.7, no.4, pp. 427 - 432.

[69].   Rama M & Syama Sundar B 2013, 'The Comparison of Antioxidative And Antimicrobial Properties of Leaf Extracts of Ocimum Gratissimum, Pimenta Dioica And Piper Betel', International Journal of Chemical And Pharmaceutical Research, Vol.2, no.1, pp.1-14.

[70].   Savita N. Ghaiwat 2014, 'Detection and Classification of Plant Leaf Diseases Using Image processing Techniques: A Review' Vol.2, no.3, pp. 2347 - 2812.

[71].   Safikul Alam.SK, Sourabh Chandra & Debabrata Samanta 2014, 'Automated Damaged Ginkgo Leaf Detection', International Journal of Advanced Research in Computer and Communication Engineering, Vol.3, no.1, pp. 5233 - 5236.

[72].   Sabah Bashir & Navdeep Sharma 2012,'Remote Area Plant Disease Detection Using Image Processing', IOSR Journal of Electronics and Communication Engineering, Vol.1, no.6, pp. 31-34.

[73].   Sanjeev S Sannakki, Vijay S Rajpurohit, V B Nargund, Arun Kumar R & Prema S Yallur 2011,'Leaf Disease Grading by Machine Vision and Fuzzy Logic', International Journal of Computer Technology and Applications, Vol. 2, no.5, pp. 1709-1716.

[74].   Sanjay B. Patil & Shrikant K. Bodhe 2011,'Leaf Disease Severity measurement Using Image Processing', International Journal of Engineering and Technology, Vol.3, no.5, pp. 297-301.

[75].   Sitaram Longani, Prof.V.V. Dixit 2013, 'Pest Detection on Leaves Using Poission's Thresholding Techniques', The International Journal Of Engineering And Science, Vol.3, no.10, pp. 6-10.

[76].   Sumathi.C.S & A. V. Senthil Kumar 2014, 'Enhancing Accuracy of Plant Leaf Classification Techniques', International Journal of Engineering Research and Applications, Vol.4, no.3, pp. 40 - 46.

[77].   Sanjay B. Dhaygude & Nitin P.Kumbhar 2013, 'Agricultural plant Leaf Disease Detection Using Image Processing', International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol.1, no.2, pp. 599 -602.

[78].   Shweta Tyagi & Hemant Amhia 2013, 'Image Enhancement and Analysis of Thermal Images Using Various Techniques of Image Processing', International Journal of Engineering Research and Applications, Vol.2, no.3, pp. 579 - 584.

[79].   Smita Naikwadi & Niket Amoda 2013, 'Advances In Image Processing for

[80].   Detection of plant diseases', International Journal of Application or Innovation in Engineering & Management, Vol.2, no.11, pp. 168 -175.

[81].   Shital Bankar, Ajita Dube,Pranali Kadam & Sunil Deokule 2014, 'Plant Disease Detection Techniques Using Canny Edge Detection & Color Histogram in Image Processing', International Journal of Computer Science and Information Technologies, Vol.52, no.2, pp. 1165 - 1168.

[82].   Suneetha.I & Venkateswarlu.T 2012, 'Enhancement Techniques for Gray scale Images in Spatial Domain', International Journal of Emerging Technology and Advanced Engineering, Vol.2, no.4, pp. 13 - 20.

[83].   Smita Patil, Shridevi Soma & Suvarna Nandyal 2013, 'Identification of Growth Rate of Plant based on leaf features using Digital Image Processing Techniques', International Journal of Emerging Technology and Advanced Engineering, Vol.3, no.8, pp. 266 - 275.

[84].   Somasundaram Devaraj, V. R. Vijaykumar, Godwin Raj Soundrarajan, Sengottaiyan & Narmada 2013, 'Leaf Biometrics Based Karyotyping of G-Band Chromosomes' International Journal of Human Genetics, Vol.7, no.4, pp.131-138.

[85].   Sandeep Kumar.E 2012, 'Leaf Color, Area and Edge Features Based Approach For Identification af Indian Medicinal Plants', Indian Journal of Computer Science and Engineering, Vol.3, no.3, pp.436-442.

[86].   Sudheer Reddy Bandi, Varadharajan A & A Chinnasamy 2013, 'Performance Evaluation of Various Statistical Classifiers In Detecting The Diseased Citrus Leaves', International Journal of Engineering Science and Technology, Vol.5, no.2, pp.298-307.

[87].   Sulochana Wadhwani, A.K.Wadhwani, Monika Saraswat 2013, 'Classification of Breast Cancer Detection Using Artificial Neural Networks', Current Research in Engineering, Science and Technology Journals. Vol.1, no.3, pp.85-91.

[88].   J.D. Sharma, Lalita Sharma & Poonam Yadav 2007, 'Antifertility Efficacy of Piper betle Linn. (Petiole) on Female Albino Rats', Asian Journal of Experimental Sciences, Vol.21, no.1, pp. 145 -150.

[89].   B. Seetha Lakshmi & K. C. Naidu 2010, 'Comparative Morphoanatomy of Piper betle L. cultivars in India', Annals of Biological Research, Vol.1, no.2, pp. 128 -134.

[90].   Saleem Shahzad & Amer Zareen 1999, 'Leaf spot of Betelvine in Pakistan, Pakistan Journal of Botany, Vol.31, no.2, pp. 437 -443.

[91]. B. C. Sarker, S. K. Adhikary, S. Sultana, A. Biswas & S. F. D. Azad 2013, 'Influence of pH on Growth and Sclerotia Formation of Sclerotium rolfsii Causal Agent of Foot Rot Disease of Betel Vine', IOSR Journal of Agriculture and Veterinary Science, Vol.4, no.1, pp. 67 -70.

[92]. A. Tripathi, R. Singh, S. K. Raj, A. P. Singh & J. K. Johri 2003, 'Molecular Identification of Phytophthora Nicotianae Isolates Causing Leaf Rot of Betelvine (Piper Betle L.)', Current Science, Vol.84, no.1, pp. 22 - 24.

[93]. Viraj A. Gulhane & Ajay A. Gurjar 2011, 'Detection of Diseases on Cotton Leaves and Its Possible Diagnosis', International Journal of Image Processing, Vol.5, no.5, pp. 590 -598.

[94]. Vinod Kumar & Rahul Raj Choudhary 2012, 'A Comparative Analysis of Image Contrast Enhancement Techniques based on Histogram Equalization for Gray Scale Static Images', International Journal of Computer Applications, Vol.45, no.21, pp.11-15.

[95]. J.Vijayakumar & Dr.S.Arumugam 2012, 'Study of Betelvine Plants Diseases and Methods of Disease Identification using Digital Image Processing', European Journal of Scientific Research, Vol.70, no.2, pp.240-244.

[96]. J.Vijayakumar & Dr.S.Arumugam 2012, 'Recognition of Powdery Mildew Disease for Betelvine Plants Using Digital Image Processing' International Journal of Distributed and Parallel Systems, Vol.3, no.2, pp.231-241.

[97]. J.Vijayakumar & Dr.S.Arumugam 2011, 'Foot Rot Disease Identification for the Betelvine Plants using Digital Image Processing', Journal of Computing, Vol. 3, no 2, pp.180-183.

[98]. J.Vijayakumar & Dr.S.Arumugam 2012, 'Early detection of powdery mildew disease for Betelvine plants using digital image analysis', International Journal of Modern Engineering Research, Vol.2, no.4, pp. 2581-2583.

[99]. J.Vijayakumar & Dr.S.Arumugam 2012, 'Powdery Mildew Disease Identification for Vellaikodi Variety of Betelvine Plants using Digital Image Processing', European Journal of Scientific Research, Vol.88, no.3, pp.409-415.

[100]. J.Vijayakumar & Dr.S.Arumugam 2012, 'Foot Rot Disease Identification for Vellaikodi Variety of Betelvine Plants using Digital Image Processing', ICTACT Journal on Image And Video Processing, Vol.3, no.2, pp.495-501.

[101]. J.Vijayakumar & Dr.S.Arumugam 2013, 'Disease Identification in Pachaikodi Variety of Betelvine Plant Using Digital Imaging Techniques', Archives Des Sciences, Vol.66, no.1, pp.308-312.

[102]. J.Vijayakumar & Dr.S.Arumugam 2013, 'Foot Rot Disease Identification for Karpoori variety of Betelvine Plants Using Digital Imaging Technique' Australian Journal of Basic and Applied Sciences, Vol.7, no.11, pp.270-274.

[103]. J.Vijayakumar & Dr.S.Arumugam 2014, 'Odium Piperis Fungus Identification for Piper Betel Plants Using Digital Image Processing', Journal of Theoretical and Applied Information Technology, Vol.60, no.2, pp.423-427.

Models and hardware implementation of methods of Pre-processing Images based on the Cellular Automata

Stepan Bilan

Faculty "Infrastructure and rolling stock of railways", State Economy and Technology University of Transport, Kiev, Ukraine

bstepan@ukr.net

ABSTRACT

The paper deals with the organization and construction of cellular automata for the implementation of the basic operations of the pre-processing images. The methods of edge detection, zoom, filling inside area of images and also selection of objects are considered. The analysis of the impact of different forms of the neighboring cells for the effective execution of operations is carried. Programs that simulate the operation of CA are developed. Computer models of the main elements in CAD Active-HDL have been obtained by modeling the structure of the CA. The obtained models have passed the test and their analysis showed high reliability of operation. This allows us to implement them in modern CPLD and FPGA hardware. This hardware is easily reprogrammed under the given structure of CA. Implementing FPGAs allows us to use one chip for realization of the basic functions of the CA. The experimental results showed that the used methods and CA are highly effective. The use of CA allows to describe of the image with high speed highly effective.

Keywords: Cellular automata, image, contour, zoom, filling inside area.

1   Introduction

All existing methods and tools for image recognition are divided into two main stages: stages of learning and recognition. Both processes are tied to the need of efficient representation and are optimal for forming an image. For advanced information technologies and computational tools the signals and images are presented in the form of codes, which values describe as an image. Therefore there is a need exists for effective methods and means, which describe an input image. Due to the variety of images and the complexity of geometric shapes and the parameters, different sets of features are used. These sets divide all images into classes. These sets of characteristic features are created by the developers and consist of different amounts.

However, each characteristical feature of different nature and requires specialized methods and tools to determine its quantitative values. In this case, for the selection of characteristical features, a great number of operations are used. The main ones are the operations of edge detection, scaling, noise removal, segmentation, selection of the objects, etc. For realization of such operations various methods and tools are used, as well as researches of new modern approaches for their effective implementation are conducted. Improvement of the performance of such operations allow cellular automata (CA), which have a unified structure, to reduce the time of operations [1-6].

2   Detection of Images Edges in CA

Edge detection operation is one of the main from image preprocessing operations. Edges analysis allows to describe of geometric shape of image of the selected object in the visual scene. Different

methods of the edges selecting are known of object, which is described in [2, 5, 7, 8, 9, 10]. However, the most effective tools that allow you to detect an edge in one cycle, are CA [2, 7]. CA allows you to implement edge detection for various organizations of cell neighborhoods (fig. 1).
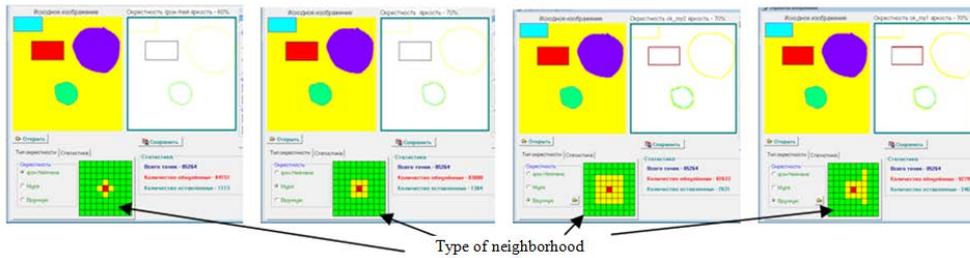

Figure 1: Example of edge detection simulation in CA

Thus, the increasing of neighborhoods order is causes by changes in thickness of the edge line. Therefore, for easing of processing of edge cells it is better to use von Neumann or Moore neighborhood. The cells function of such CA for neighborhood of the von Neumann can be represented by the following relationship

$$b(t+1) = \begin{cases} b(t), & if \quad x_1(t) \wedge x_2(t) \wedge x_3(t) \wedge x_4(t) = 0 \\ 0, & if \quad x_1(t) \wedge x_2(t) \wedge x_3(t) \wedge x_4(t) = 1 \end{cases},$$ (1)

where $x_1(t)$, $x_2(t)$, $x_3(t)$, $x_4(t)$ - the signals at the outputs of neighborhood cells (top, right, bottom, left) at time t; $b(t)$ - state of a cell at time t.

For Moore neighborhood the model (1) will contain eight signals from neighborhood cells

$$b(t+1) = \begin{cases} b(t), & if \quad x_1(t) \wedge x_2(t) \wedge ... \wedge x_8(t) = 0 \\ 0, & if \quad x_1(t) \wedge x_2(t) \wedge ... \wedge x_8(t) = 1 \end{cases}.$$ (2)

As we can see from the illustrated example it is possible to create a shadow of the object, to make some lines thicker, and the others are thinner. Through variation of the neighborhood structure we can create different shapes and transformation of the objects contours.

However, for halftone image processing and edge detection of the object it is necessary to use thresholding processing. The CA separates cells of object from the background on the base of selected threshold value of brightness. Each cell is determined its the mean value from value set of cells of the neighborhood and is compared with a predetermined threshold value. If the value exceeds the control cell threshold, this cells belongs to the cells of the edge. Otherwise, the cell is set to value of background (fig. 2).
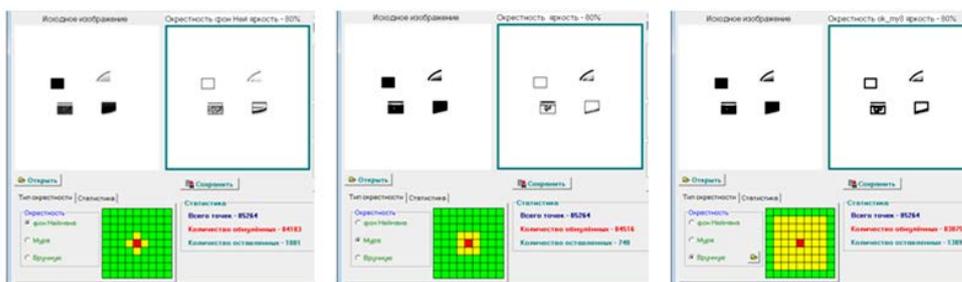
Figure 2: Example of edge detection of halftone image

Threshold variation value allows to select edge as well as to implement noise filtering, which presented in the image. Also, this approach allows the edges selection of images by gradations of brightness.

Cell structure, which implements the CA for edge detection of the halftone image is submitted on figure 3.
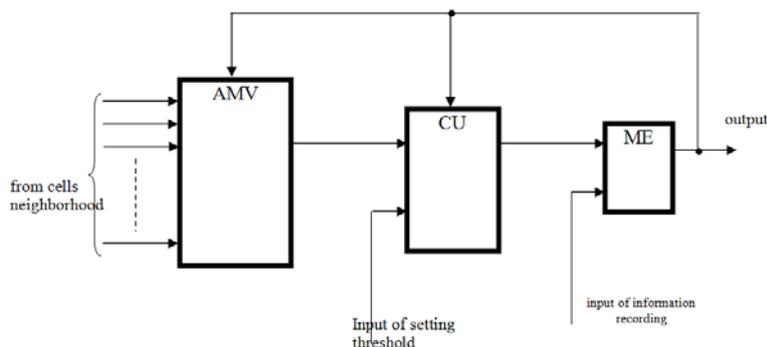

Figure 3: Structural circuit of one cells of CA for edge detection

Cell contains a memory element (ME), comparing unit (CU) and the adder of mean value (AMV) that summarizes values of the neighborhood cells and determines the average value. CU compares it with a predetermined threshold value and with code from output of the ME, and accordingly controls the state of ME.

For all the selected types of neighbors their comparative analysis was performed by number of cells that form edges.

Graph of cell contour distribution for the neighborhood of von Neumann, Moore, of the second and third orders is shown in Figure 4.
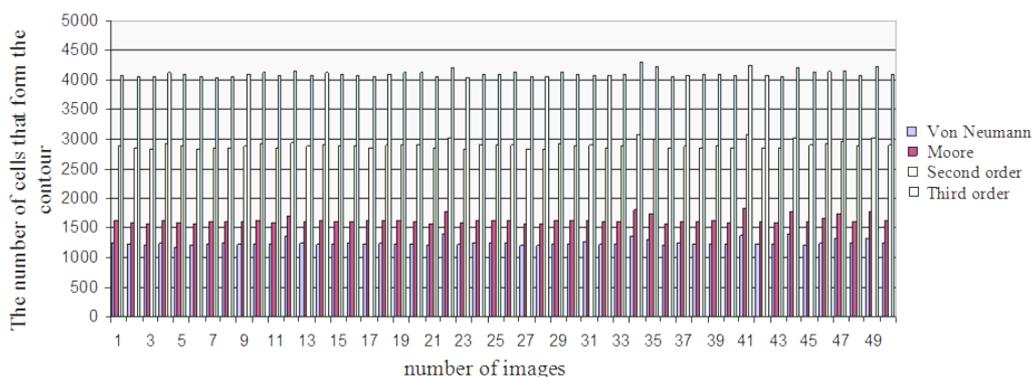
Figure 4: Graph of cell contour distribution for the neighborhood of von Neumann, Moore, of the second and third orders

The analysis of the dependency shows the stability of the difference between the numbers of cells that form the edges. At the same time shape of the curves are the same for all types of neighborhoods at the same sample of images, which have been investigated.

The experiments were conducted for symmetric structures of neighborhoods that have the same shape but different in the number of cells that form the neighborhood.

In the next part of the experimental research real images of figures obtained by photographing were used.

The resulting image was divided into separate images of the equal dimension (300×300). At the same time each figure was rotated on a certain angle.

For the contrast selection and binarization of the resulting images boundaries with effective sensitivity were defined where unneeded cells were completely removes, which did not belong to the figure. For existing an image binarization with sensitivity in the range from 50 to 60 percent was carried out. Image processing was carried out by the program "Study". Histograms of cells that form the edges of the figures images for given forms of neighborhood are shown in Figure 5.
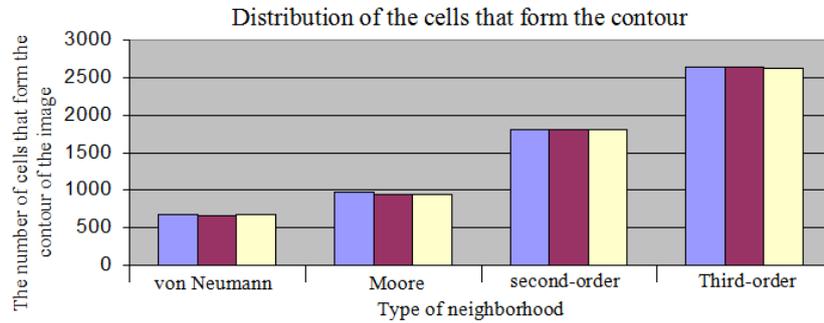
Figure 5: Histograms of edge cells of images for selected types of neighborhoods

With the help of these experiments average values of the ratios of the number of cells that form a contour for different types of neighborhoods were obtained. The relations obtained show that the difference between the obtained values of the ratio to ratio of real images and the computer images, practically negligible. Consequently, all depends on the accurate initial morphological image processing by the camera.

3   Scaling of images in CA

There are many problems in the known image scaling methods that are associated with the assignment of the center of scaling and arbitrary scaling factors for X and Y coordinates [10, 11]. CA allows to solve these problems.

The method is described below:

Initially, a cell is selected, which is the center of scaling. Each cell of CA is given by a scaling factor K_M. The value K_M is determined. If $K_M > 1$ cell reproduction begins with shift on both coordinates in turn. The closest cells to the central cell first begin to reproduce. The other cells and their values shift. The first cell creates additional cells on one of the coordinates with the same conditions. Their number is K_M. The other cells and their values shift on the K_M of cells from the cell of scaling center (figure 6).
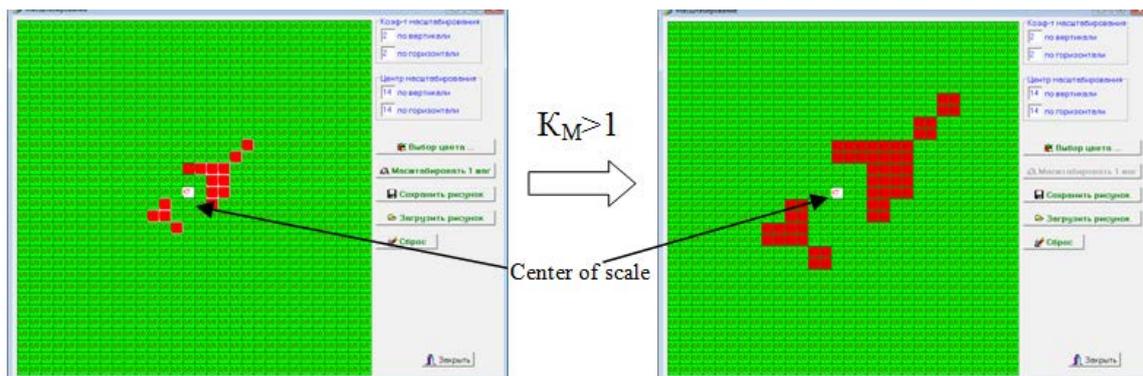

Figure 6: Example of scaling images based on CA

When all of the cells and their values on one coordinate proliferate in opposite sides from the cell of scaling center, cells reproduction begins according to another coordinate. The reproduction is carried out similarly. At different scaling factors for different coordinates also the image changes in the direction of these coordinates (see figure 7).
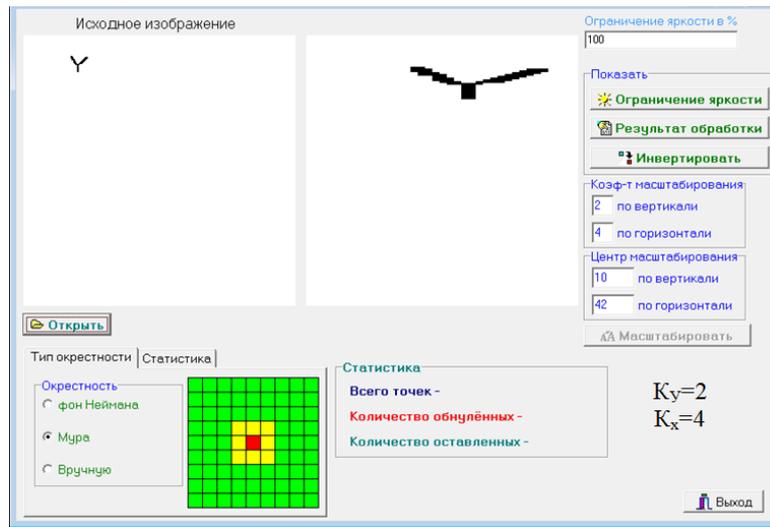

Figure 7: Example of scaling images based on CA with different KM

If $K_M < 1$, then the cells are absorbed in the direction to the cell in the center of scaling. The first absorbed cells are the neighboring cells with equal values, which are adjacent to the central cell. Their number is equal to K_M. Other cells and their values shift to the center on the value K_M of cells. After the first absorption the cells with another state begin to absorb. These cells are located after the first absorbed cells. This process continues until the complete enumeration of all the cells on one coordinate, and then on the other.

Model of the scaling process description can be divided into two models that describe the functioning of the cell for increase or to decrease of the image. Each of these models takes into account the three components.

- States of a cell at time t for one of the coordinates $b_i(t)$.
- Availability of excitation signal $e_i(t)_{excite}$.
- The scale ratio value recorded in this cell.

Model describing the cell behavior for process of the image increase on the X-axis in the one direction is the following

$$\frac{b_i^x(t+1)}{K_i^M(t+1)} = \begin{cases} \frac{b_{i-1}^x(t)}{K_i^M(t)+K_{i-1}^M(t)}, & \text{if } b_i^x(t)=b_{i-1}^x(t), \ e_{i-1}^x(t)_{excite}=1 \\[6pt] \frac{b_i^x(t)}{0}, & \text{if } b_i^x(t)=b_{i+1}^x(t), \ e_i^x(t)_{excite}=1, \\[6pt] \frac{b_i^x(t)}{0}, & \text{if } b_i^x(t)\neq b_{i+1}^x(t), \ e_i^x(t)_{excite}=1, \ s_{i+1}^x(t)_{shift}=1, \ K_i^M(t)>0 \\[6pt] \frac{b_{i-1}^x(t)}{K_i^M(t)-1}, & \text{if } b_{i-1}^x(t)\neq b_i^x(t), \ e_{i-1}^x(t)_{excite}=0, \ K_i^M(t)>0, \ s_{i+1}^x(t)_{shift}=1, \\[6pt] \frac{b_{i-1}^x(t)}{K_{i-1}^M(t)}, & \text{if } e_{i-1}^x(t)_{excite}=0, \ K_{i-1}^M(t)=0, \ s_{i-1}^x(t)_{shift}=s_{i+1}^x(t)_{shift}=1 \\[6pt] \frac{b_{i-1}^x(t)}{K_{i-1}^M(t)}, & \text{if } b_{i+1}^x - \text{does not exist}, \ e_{i-1}^x(t)_{excite}=e_i^x(t)_{excite}=0, \ s_{i-1}^x(t)_{shift}=1 \\[6pt] \frac{b_{i-1}^x(t)}{0}, & \text{if } b_{i+1}^x - \text{does not exist}, \ e_{i-1}^x(t)_{excite}=1, \ K_{i-1}^M(t)>0 \end{cases} \tag{3}$$

where $\dfrac{b_i^x(t+1)}{K_i^M(t+1)}$ - the status value of the cell (numerator) at time (t +1) and the value of scaling factor

(the denominator) of cell, which is i-th cell on the X coordinate of the cell selected as the center of scaling;

$e_i^x(t)_{excite}$ - the value on output of formation of a excitation signal of the cell at time t, which is the i-th cell on the X coordinate of the cell selected as the center of scaling;

$s_{i+1}^x(t)_{shift}$ - Signal, which allows record values of the state of the i-th cell in the (i +1)-th cell.

In the numerator of the model internal state of the cell is described, and the denominator represents the value of the scaling factor.

The first row in (3) describes the increase in the neighboring cells $b_i^x(t)$, which have the state, and the i-th cell is selected from the center of scaling becomes excited. An example of such a situation is shown in Figure. 8.
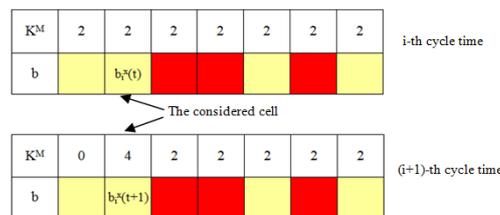
Figure 8: Example of the cell functioning in a situation as described in the first row of model (3)

Second row of model (3) describes the transmission of the excitation signal next adjacent cell (Fig. 9), which is in the same state as the cell excitation.
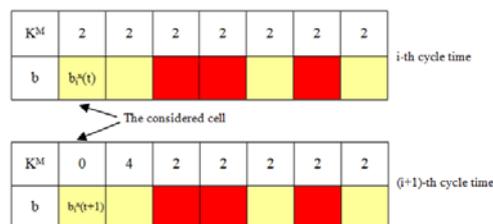
Figure 9: Example of excitation signal transmission

The third row of the model (3) shows that the i-th cell records its state in (i +1)-th cell (Fig. 10).
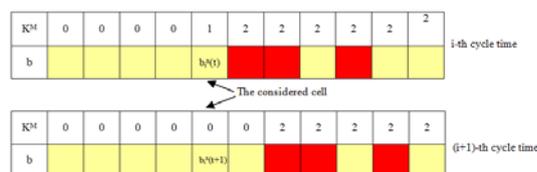
Figure 10: Example implementation of the model (3) according to an embodiment of the third row

The fourth row describes the transition of state of i-th cell into the state (i-1)-th cell. State of the i-th cell is transferred to the (i +1)-th cell (Fig. 11).
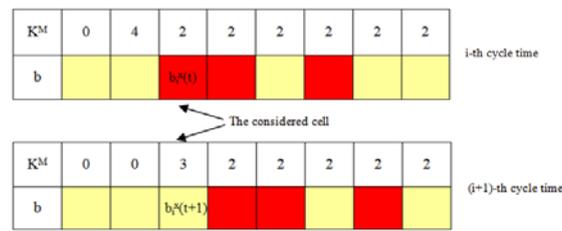
Figure 11: Implementation example of the fourth row in model (3)

The fifth line describes a behavior cell in case of shift of its state from cell center of scaling (Fig. 12).
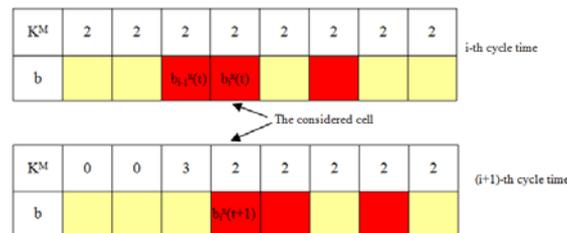

Figure 12: An example of transmitting state of the cell, which is not excited

The sixth and seventh rows describe the behavior of the boundary cells of CA.

At the initial time every cell records a value K^M, which decremented by 1 ($K^M - 1$).

Thus, the excited state of the cells at each time step is transferred to the next neighboring cells, if $K^M(t) > 0$. The number of such cells is equal K^M.

The cell states are recorded in the following neighboring cells, if they received shift signals from two neighboring cells (shift of cell states in the direction from the cell, which is the center of scaling). Installing cells in readiness to shift is carried out by distribution shift signal from the center cell of scaling. The direction of propagation of this signal indicates the direction of the shift.

The excitation signal is generated at the next adjacent cell based on a specified condition

$$e_i(t+1)_{excite} = \begin{cases} 1, & if \quad K_{i-1}^M(t) = 0, \quad e_{i-1}(t)_{excite} = 1 \\ 1, & if \quad b_i(t) = b_{i-1}(t), \quad e_{i-1}(t)_{excite} = 1 \\ 0, & in \quad other \quad cases \end{cases}$$ (4)

To zoom out, of an image the absorption cell states are used. Cells are absorbed in each row and by each coordinate. Initially are absorbed cells along the X-coordinate and then along the Y coordinate. The model describing the functioning of the cells in direction X coordinate has the form

$$\frac{b_i^x(t+1)}{K_i^M(t+1)} = \begin{cases} \dfrac{b_i^x(t)}{K_i^M - 1}, & if \quad e_i^x(t)_{excite} = 1, \ b_i^x(t) = b_{i+1}^x(t), \\ \dfrac{b_{i+1}^x(t)}{K_{i+1}^M(t)}, & if \quad e_{i-1}^x(t)_{excite} = 1, \ K_{i-1}^M(t) > 0, \ b_i^x(t) \neq b_{i+1}^x(t), \ b_i^x(t) = b_{i-1}^x(t), \ s_i^x(t)_{shift} = 0, \\ \dfrac{b_{i+1}^x(t)}{K_{i+1}^M(t)}, & if \quad e_i^x(t)_{excite} = e_{i-1}^x(t) = 0, \quad s_{i-1}^x(t)_{shift} = s_i^x(t) = 1 \\ \dfrac{b_i^x(t)}{0}, & if \quad e_i^x(t)_{excite} = 1, \ b_i^x(t) = b_i^x(t+1) \\ \dfrac{b_i^x(t)}{0}, & if \quad b_{i+1}^x(t) - \text{does not exist} \end{cases}$$ (5)

The first row describes the behavior of cells in the case when the next adjacent cell from the excited cells has the same state (Fig. 13).

Figure 13: Example of the cell functioning which describes in the first row of the model (5)

The second row describes the shift of cells in row in direction the neighboring excited cells (Fig. 14).
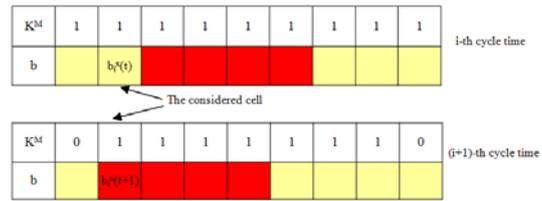

Figure 14: Example of the shift of cell states

The third row of model (5) describes the shift of the contents of all cells in the direction of the excited cells. In this case, the cells are not have excitation cell among adjacent cells (Figure 15).
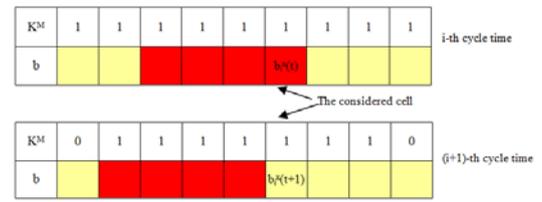
Figure 15: Example of the shift excited states of a cell without adjacent excitation cells

The fourth row describes the process of absorption by one cell its adjacent cells with equal state (Figure 16).
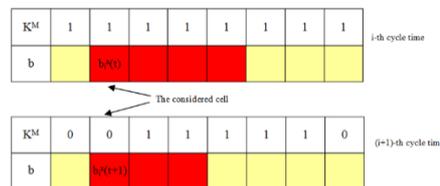

Figure 16: Example of absorption of the cell

If state of the cell is shifted or absorbed, then this cell generates the shift signal $s(t)_{shift}$ to neighboring cells.

The excitation signal is generated at the next adjacent cell based on a specified condition

$$e_i(t+1)_{excite} = \begin{cases} 1, & if \quad K_{i-1}^M(t) = 0, \quad e_{i-1}(t)_{excite} = 1, \quad b_{i-1}(t) = b_i(t) \\ 1, & if \quad b_{i-1}(t) \neq b_i(t), \quad e_{i-1}(t)_{excite} = 1 \\ 0, & in \quad other \quad cases \end{cases} \quad , \quad (6)$$

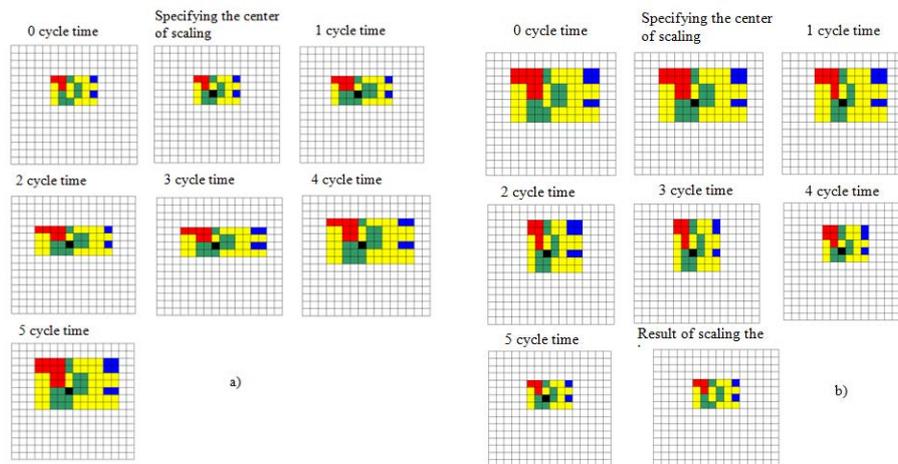An example scaling operation on every cycle time is represented by Figure 17.

Figure 17: Example of a detailed image scaling: a – scaling with $K^M > 1$, b – scaling with $K^M < 1$

According to the proposed methods a program was created that implements them on the basis of models of CA functioning. Examples of scaling of real images by program are shown in Figure 18.
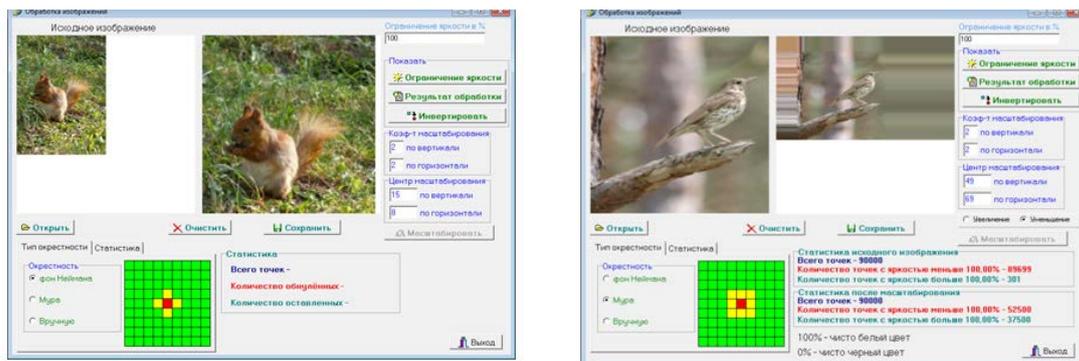

Figure 18: Example of scaling images based on CA

For accurate and reliable research of scaling figures were formed, which consisted of units of cells. Small number of cells that forms the image of figures was used for the precise counting them during zooming.

At the time of the research of cells scaling algorithm was carried out by counting the number of cells that form the image after applications of each zoom level. Histogram of cells distribution on the images of figures with equal amounts of cells, which form the primary image is shown in Figure 19.

Obtained dependence shows that the number of cells, which is obtained when the zoom factor is increased on 1, increases in four times, and when the zoom factor is 3, the number of cells increases in 16. The dependence shows an increase of cells numbers in four times more then K_M.
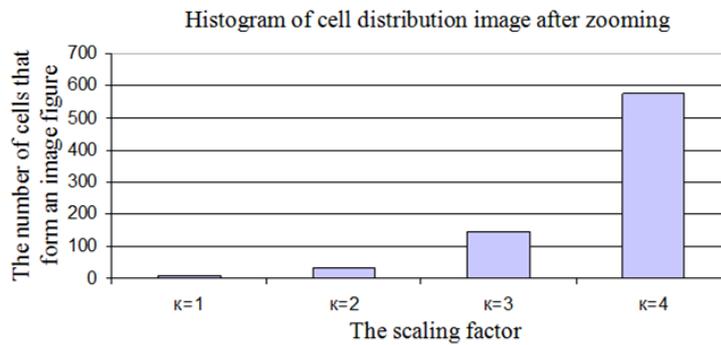
Figure 19: Histograms of cells distribution contour of figures for specific types of neighborhood

4   Filling in CA

Reverse operation for edge detection is operation of filling of inside area of image boundary. These operations have differences under time expenses. If edges have been selected during one clock cycle in the CA, time of operation of area filling depends on the number of CA cells that make up the inner area of the contour.

Known methods of parity control and seed cells are successfully implemented in the CA. However, the implementation of parity control method complicates the structure of the cells due to the need to eliminate false results at the presence of extreme peaks and horizontal lines (fig. 20). The implementation of the method of the seed pixel cells simplifies the cell structure and requires to define the inner area (fig. 21).
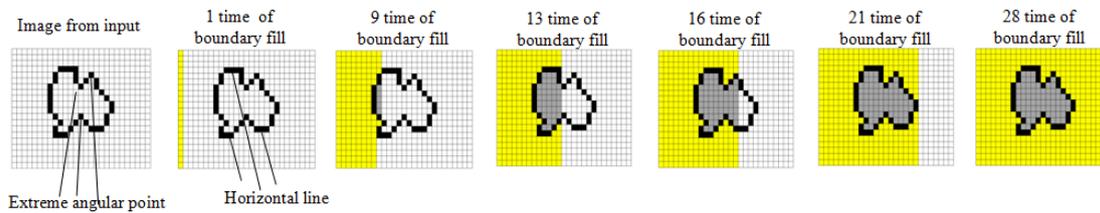

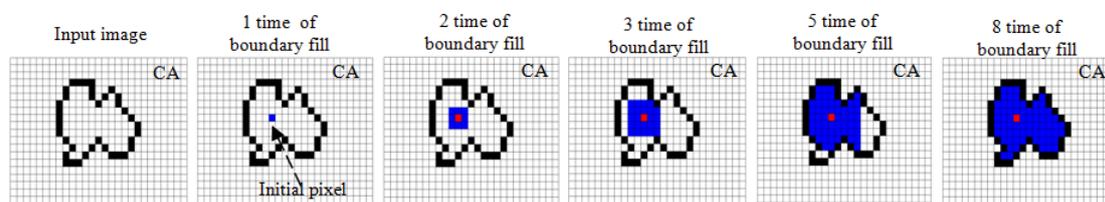Figure 20: Example of implementation of the parity control method

Figure 21: Example of implementation of the parity control method

Example of implementation of the seed cells method.

To implement the parity control method the following model is used.

$$b_i(t+1) = \begin{cases} b_{i-1}(t), & if \quad b_i(t) = a(t)_{backgr}, \quad e_i(t)_{excite} = 1, \quad b_{i-1}(t) \neq a(t)_{backgr}, \quad F_i(t)_{extreme} = 1, \quad X_i(t)_{even} = 1, \\ b_i(t), & in \quad othe \quad cases \end{cases}$$ (7)

where

$$F_i(t)_{extreme} = \left[x_i(t)_{L.UP} \vee x_i(t)_{C.UP} \vee x_i(t)_{R.UP} \right] \wedge \left[x_i(t)_{L.B} \vee x_i(t)_{C.B} \vee x_i(t)_{R.B} \right],$$ (8)

$$X_i(t+1)_{even} = \overline{X_{i-2}(t)} \wedge X_{i-1}(t) \wedge F_{i-1}(t)_{extreme} \wedge \overline{X_i(t)} \vee X_{i-2}(t) \wedge X_{i-1}(t) \wedge$$
$$\wedge \overline{F_{i-2}(t)_{extreme}} \wedge \overline{X_i(t)} \wedge F_i(t)_{extreme}$$

(9)

$a(t)_{backgr}$ - the value of the background brightness;

$X_i(t)$ - the output signal of the i-th cell at time t;

$x_i(t)_{even}$ - the value of the signal of paired values of edge in the i-th cell one row CA (1 − is odd intersection of contour, 0 - a multiple of two);

$x_i(t)_{L.UP}, x_i(t)_{C.UP}, x_i(t)_{R.UP}$ - the signals states of cells that present left, center and right upper cells of the Moore neighborhood for a given cell at time t;

$x_i(t)_{L.B}, x_i(t)_{C.B}, x_i(t)_{R.B}$ - the signals states of cells that present left, center and right bottom cells of the Moore neighborhood for a given cell at time t.

The model describes the operation of the single cell of CA in row. The signals at the outputs of the neighborhood cells and their state at the appropriate time are taken into account.

For filling of the image by seed cell method we uses the model

$$b_i(t+1)_{excit} = \begin{cases} 1, & if \quad Q_i(t)_{neugh}^{excit} = 1, \quad b_i(t) = a(t)_{backgr} \\ b_i(t), & if \quad Q_i(t)_{neugh}^{excit} = 0, \\ b_i(t) & b_i(t) = 1, \quad Q_i(t)_{neugh}^{excit} = 1 \end{cases}$$

(10)

where $Q_i(t)_{neugh}^{excit} = \bigvee\limits_{j=1}^{n} b_j(t)_{excit}$ - indicates the presence of the excited cells in the neighborhood of the i-th cell, which has the state 1;

$b_j(t)_{excit}$ - the output signal of j-th cell of neighborhood, which is in an excited state;

n – the number of cells, which form a neighborhood;

$b_i(t+1)_{excit}$ - the state of the cell, which is in an excited state at time t+1.

The values originally recorded in cells must be fixed and have a certain state after filling. Considering this, the cell has the structure shown in Figure 22.
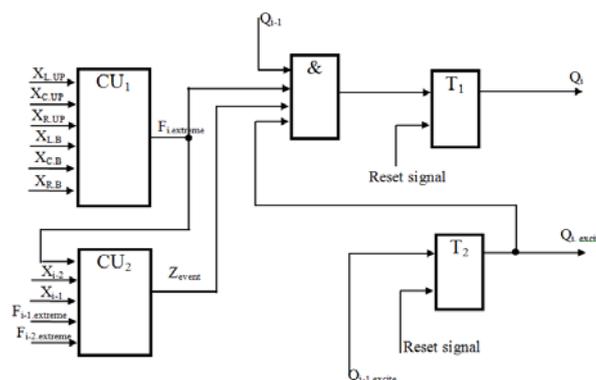

Figure 22: Scheme of cell, which eliminates situations of a false filling of areas that do not belong to the image

T_1 is intended for recording an image, and T_2 – for supporting of the filling process (trigger of excitation). Control units (CU_1 and CU_2) are provide control of T_1 trigger state.

On the scheme the output status i-th cell is shown Q_i, and the output of the excitation signal - $Q_{i.excite}$.

The main condition is that in the contour of image should not be any discontinuities, because this leads to errors in the implementation of the method. Organization of cells of the medium is based on the two triggers, one of which is used to store information about the boundary, and the second - to fill of area.

For a hardware implementation of the method of seed cells, each cell has additional trigger (T_c) of the initial installation. In the state "1", the edge cells will be initially. The speed of the filling process of an image depends on the right choice of cell. Filling time is also reduced with an increasing number of seed cells.

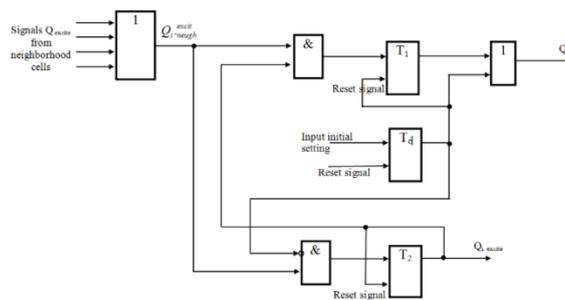Functional scheme of the cell is shown in Figure 23.

Figure 23: Functional scheme of cell to implement the method of the seed cells

Implementation of these methods in CA absentia reduces time costs; their conjugation allows halftone filling based on the location of the radiation source.

The second method is faster than the first. The speed increase allows using of several initial image elements which fill the inner area. However, the initial cell is mounted inside the area which is defined by the user in advance. This is a drawback of the second method.

5 System Structure of Image Recognition on the Base of CA

The considered operation of images pretreatment are ones of the basic operations, which allow to simplify the methods of selection of the information characteristic features. Use of CA allows to combine the group of CA for all operations in one CA. The structure of cell of generalized CA is shown in Figure 24.
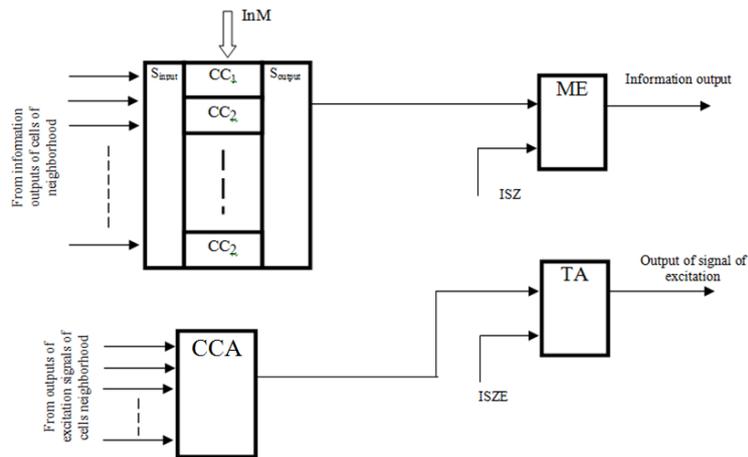
Figure 24: Functional scheme of cell of generalized CA

One cell of a generalized CA consists of memory element (ME) and control scheme of its condition (CSC). CSC consists of n combinational circuits (CC), which are combined on the common inputs by switcher inputs (S_{input}) from the outputs of cells that belong to the neighborhood. The outputs of each CC are connected to the inputs of the switch outputs (Soutput). Its total output is connected to the control input of the state of ME. Through the control input (InM) selects the CC, which must operate at a given time.

Each CC has a structure, which realizes a necessary operation (edges selection, scaling, filling, etc.), and its output at a predetermined timing is connected to a common output. Signal from the output CSC controls the state of ME.

This structure allows to simplify CA and by switching of CC to reach different sequences of operations. To extending the number of operations performed by the CA an additional trigger of arousal (TA) and an appropriate combinational circuit (CCA), which processes the signal from output of the TA of neighborhood cell are introduced in each cell. Availability of TA allowed to implement an operation specified by cells, where TA are in a single state. The excitation signal is transmitted between neighboring cells. Introduction of additional TA allows to divide the processes of performing of different operations to prepare images. Cell structure for CA, which implements the described operations, were modeled in CAD Active-HDL and showed reliable operation. The resulting model can be easily implemented in modern FPGAs.

After performing operations of previous selection the operation of the main characteristic features definition is conducted and a vector of numerical values of attributes is formed. This vector describes the image on input of recognition system. The requirements are imposing as to the location and quantitative characteristics. One of the approaches to description of the images on the digital and the analog level is described in the works [2, 12]. This method implements the functioning of the system as an analogue of the human optic canal at the level of the waveform.

Block - circuit of the system processing and recognition is presented in Figure 25.
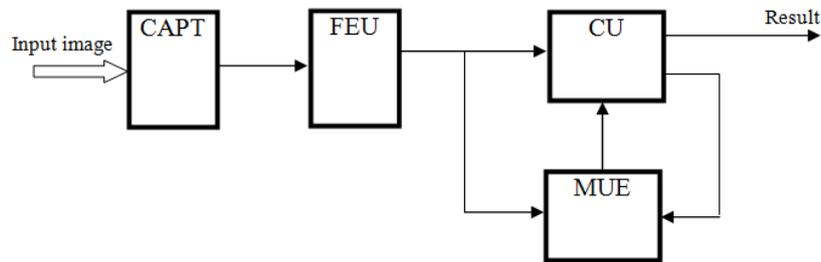
Figure 25: Structure of system of processing and recognition of image

The system comprises CA of pretreatment (CAPT), feature extraction unit (FEU), the memory unit of etalons (MUE) and a comparison unit (CU). CAPT processing the input image and FEU selects the characteristic features. At the output FEU vector of numeric values of attributes, which describe the image is formed. One of the variants of FEU is presented in the works of [2, 12, 13]. The code of vector is supplied to a first input of the CU, which performs the comparison with the etalon code vector. Etalon vector is supplied to the second input of the CU from the output MUE. If the code of etalon does not match to the code of input image, the CU generates a signal on the second output, which is fed to the second input of MUE. MUE carries out is forming of the next closest code of etalon, etc. If the code is not found, the code of the input image is assigned an identifier and it is stored in the MUE through the first input. The system recognizes the image and learns during functioning.

6   Conclusion

All operations are implemented on the CA, which implement asynchronous signal propagation of excitation in the field of CA. The excited cells perform the selected operation and conduct an analysis of cells of their own neighborhood. This approach made possible to perform operations in the individual cells of a field of CA and expanded the number of different operations, which are performed simultaneously.

As a result of the executed experiments on the basis of computer simulation high efficiency of use CA is proved. The developed tools allow to specify the shape of the surrounding area and the thickness of the contour, and to choose an arbitrary scaling center. Besides the method allows to specify different coefficients for each coordinate, which allows to set the appropriate image transformation. By the individual customization of CA efficient extracting of the necessary features performed, which enable optimal description of both simple and complex images.

REFERENCES

[1].   Von Neumann, J. Theory of Self-Reproducing Automata: Edited and completed by A. Burks. University of Illinois Press, 1966.

[2].   Belan S., Specialized cellular structures for image contour analysis, Cybernetics and Systems Analysis, 2011. 47(5): p. 695-704.

[3].   Wolfram S., Cellular Automata. Los Alamos Science, 1983. 9: p. 2-21.

[4].   Belan S., and Motornyuk R., Extraction of characteristic features of images with the help of the radon transform and its hardware implementation in terms of cellular automata. Cybernetics and Systems Analysis, 2013. 49(1): p. 7-14.

[5].   Ioannidis K., Andreadis I. Sirakoulis G. Ch. An Edge Preserving Image Resizing Method Based on Cellular Automata. Springer-Verlag, Berlin Heidelberg, ACRI2012, LNCS, 749, 2012. 375-384.

[6].   Bandini, S., Bonomi, A., Vizzari, G. An Analysis of Different Types and Effects of Asynchronicity in Cellular Automata update Schemes. Natural Computing, 2012. 11(2): p. 277-287.

[7].   Kozhemyako V., Belan S., Savaliuk I. (1997). Optoelectronic self – reguletion neural system for treatment of vision information. SPIE Proceedings, Wasington, USA, 1997. 3055, p. 120-126.

[8].   Chen M.J., Huang C.H., Lee W.L. (2005). A fast edge-oriented algorithm for image interpolation. Image and Vision Computing, 2005. (23): p. 791–798.

[9].   Canny J. A computational approach to edge-detection. IEEE Trans. Pattern Anal. Mach. Intell, 1986. (8), p. 679–700.

[10]. Wolfram S. Theory and applications of Cellular Automata. World Scientific, Singapore, 1986.

[11]. Amanatiadis A., Andreadis I., Gasteratos, A. A Log-Polar interpolation applied to image scaling. In: IEEE International Workshop on Imaging Systems and Techniques, Cracovia, Poland, 2007, p. 1–5.

[12]. Cha Y., Kim S. The error-amended sharp edge (EASE) scheme for imaging zooming. IEEE Trans. Image Process, 2007. (16), p. 1496–1505.

[13]. Belan S., Belan N. Temporal-Impulse Description of Complex Image Based on Cellular Automata. Springer-Verlag Berlin Heidelberg. PaCT2013, LNCS.- 2013. 7979. 291-295.