ICASC

1.Interdisciplinary Conference
on Applied Soft Computing

March 3–4, 2008

PROCEEDINGS

About the conference

The ICASC aim to offer a presentation and networking platform for people who apply soft computing techniques (e.g. neural networks, genetic algorithms, genetic programming, agent simulations, fuzzy systems) to solve computational problems that are either concerned with real life problems or with scientific questions about the organisation of complex structures (like the organisation of natural, emergent or artifical networks).

The ICASC will be held in Münster, Germany on **March 03-04, 2008**.

Important Dates

Deadline: **15 January 2008**
Notification: **31 January 2008**
Conference: **3-4 March, 2008**

Scientific board

The conference will be organized by the

Faculty of Mathematics and Natural Science
Department of Mathematics and Computer Science
Institute for Computer Science

in cooperation with the

Institute of Neuro and Behavioural Biology,
Dept. of Experimental Tumorbiology

and the

Research Group No. 4
of the Interdisciplinary Center for Clinical Research
Klinik and Poliklinik for Psychiatry und Psychotherapy

Chairmen

Sascha Hauke
Dominik Heider
Martin Pyka

Conference site

The conference will take place in the

Institute of Neuro and Behavioural Biology
Badestraße 9
48149 Münster, Germany

More information at
http://www.muenster.de/

How to arrive

By bus or train: Upon arriving at Münster (Westf.) Hauptbahnhof (central train station), take city bus lines 11, 12, 13 or 22 to bus stop “Landgericht”. Cross at the traffic light slightly up the street. You are now at the intersection with Baadestrasse. Enter Baadestraße, the institute will be on your left hand side.

By car: From the south, take the A1 to highway junction Kreuz Münster-Süd. Follow the signs to Münster “Centrum” via B51 and Weseler Straße. After passing the Aasee on your left, continue on straight on Am Stadtgraben for another 500 meters, then turn left onto Gerichtstraße. After 200 meters turn left into Baadestraße, the institute will be on your left hand side. From the north, take the A1 to highway junction Kreuz Münster-Nord, continuing on the B54/Steinfurter Straße towards Münster “Centrum”. After passing the “Schloß” (palace) on your right hand side, turn right onto Gerichtstraße. After 200 meters turn left into Baadestraße, the institute will be on your left hand side.

Programme

	Monday	Tuesday
10:00:00	**Opening**	
10:15:00	**Keynote – Markus Borschbach** *University of Muenster, Germany*	
11:20:00	**Martin Pyka** *University of Muenster, Germany*	**Prabavathy Kethsy** *Anna University, India*
12:00:00	**Lunch**	**Lunch**
13:30:00	**Dominik Heider** *University of Muenster, Germany*	**M.N Kavitha** *Anna University, India*
14:00:00	**Sascha Hauke** *University of New Brunswick, Canada*	**Bastani Dariush** *Sharif University of Technology, Iran*
14:30:00	**Amrit Kaur** *Punjab Technical University, India*	**Mohammad Saleh Shafeeyan** *University of Teheran, Iran*
15:00:00	**Coffee break**	**Coffee break**
15:30:00	**Wei Zhang** *Wuhan University, China*	**De Indrajit** *MCKV Institute of Engineering, India*
16:00:00	**Devi Nirmala** *Anna University, India*	**Organization Discussion Panel**
16:30:00	**Scientific Discussion Panel**	**Paper Awards & Closing Remarks**
20:00:00	**Social Event**	

SCALABLE CHARACTER SKELETON USING WAVELET METHOD

M.N.Kavitha
II M.E CSE
Department of Computer Science
and Engineering
K.S.Rangasamy College of Technology,
KSR Kalvi Nagar,
Thiruchengode-637 209
Tamil Nadu, India.
91-9843594196
kavithafeb1@yahoo.co.in

ABSTRACT

The development of proposed model was inspired by the polygonal line algorithm and is extended to find principal graphs and complemented with two steps specific to the task of skeletonization i.e., an initialization method to capture the approximate topology of the character, and a collection of restructuring operations to improve the structural quality of the skeleton produced by the initialization method. Character skeleton plays a significant role in character recognition. The strokes of a character may consist of two regions, namely, singular and regular regions. The intersections and junctions of the strokes belong to singular region, while the straight and smooth parts of the strokes are categorized to regular region. Therefore, a skeletonization method requires two different processes to treat the skeletons in theses two different regions. To overcome the problems of complex computation, locating central line of the stroke and distortion of singular region, a new scheme of extracting the skeleton of character based on wavelet transform is presented in this thesis. It consists of extraction of primary skeleton in the regular region, amendment processing of the primary skeletons and connection of them in the singular region. The system tested the algorithm on isolated handwritten digits and images of continuous handwriting. The results indicate that the proposed algorithm finds a smooth medial axis of the great majority of a wide variety of character templates and substantially improves the pixel-wise skeleton obtained by traditional thinning methods. This scheme is applicable to not only binary image but also gray-level image.

Categories and Subject Descriptors

I.4.10 Image Representation [Morphological]

General Terms

Algorithms, Documentation, Experimentation.

Keywords

Skeletonization, wavelet transform, singular region, regular region.

1. INTRODUCTION

Skeletonization is one of the important areas in image processing. It is most often, although not exclusively, used for images of handwritten or printed characters in this context. When look at the image of a letter, see it as a collection of curves rather than a raster of pixels. Since the earliest days of computers, it has been one of the challenges for researchers working in the area of pattern recognition to imitate this ability of the human mind.

Approaching skeletonization from a practical point of view, representing a character by a set of thin curves rather than by a raster of pixels is useful for reducing the storage space and processing time of the character image. It was found that this representation is particularly effective in finding relevant features of the character for optical character recognition.

The objective of skeletonization is to find the medial axis of a character. Ideally, the medial axis is defined as a smooth curve (or set of curves) that follows the shape of a character equidistantly from its contours. In case of handwritten characters, one can also define the medial axis as the trajectory of the pen stroke that created the letter. Most skeletonization algorithms approximate the medial axis by a unit-width binary image.

In one of the most widely-used strategies, this binary image is obtained from the original character by iteratively peeling its contour pixels until there remains no more removable pixel. The process is called thinning and the result is the skeleton of the character. The different thinning methods are characterized by the rules that govern the deletion of black pixels.

The development of the method was inspired by the apparent similarity between the definition of principal curves

and the medial axis. Principal curves were defined as "self-consistent" smooth curves which pass through the "middle" of a d-dimensional probability distribution or data cloud, whereas the medial axis is a set of smooth curves that go equidistantly from the contours of a character. Therefore, by representing the black pixels of a character by a two dimensional data set, one can use the principal curve of the data set to approximate the medial axis of the character.

The basic idea of the polygonal line (PL) algorithm is to start with a straight line segment and, in each iteration of the algorithm, increase the number of segments by one, by adding a new vertex to the polygonal curve produced in the previous iteration. After adding a new vertex, the positions of all vertices are updated in an inner loop so that the resulting curve minimizes a penalized distance function.

Generally, the skeleton of a character is the locus of the midpoints or the symmetric axis of the character stroke. Different techniques of the local symmetry analysis may generate different symmetric axes (skeletons). A great deal of skeletonization techniques have been proposed, among which there are three well-known algorithms of the local symmetry analysis: the symmetric axis transform (SAT), smoothed local symmetry (SLS), and process-inferring symmetry analysis (PISA).

2. LITERATURE REVIEW

Simon [7] partitioned the character stroke into regular and singular regions. The singular region corresponds to ends, intersections and turns, and the regular region covers the other parts of the strokes. Thus, the skeletonization process of character based on the symmetry analysis contains two main steps:

1) Computing the primary skeleton in the regular region and

2) Amendment processing the skeleton in the singular region.

An essential challenge for computing the primary skeleton is that it is difficult to decide the symmetric pairs from the boundaries of the character strokes in both continuous and discrete domains [8]. Particularly, in the discrete domain, it may not be possible from one side to find the symmetric counterparts on the opposite side of the contour such that they form the local symmetries. Even if a pair of symmetric contour pixels can be exactly found, their corresponding symmetric center may not exist exactly.

The technique of the regularity-singularity analysis, which is based on the constrained Delaunay triangulation and nonsymmetrical approach, is a sound solution for this problem [4, 9, and 10]. Unfortunately, its implementation often suffers from the complicated computation, as well as the three traditional symmetry analysis [5, 6]. Moreover, the nonsymmetrical approaches produce the skeleton with strong deviation from the center of the underlying character strokes and rough human perceptions [10].

Character strokes contain a variety of intersections and junctions, which belong to the singular region. For such cases, an other major problem of the traditional methods is that their results do not always conform to the human perceptions, since they often contain unwanted artifacts such as noisy spurs and spurious short branch in the singular region [1, 2, and 3]. To compute the skeleton in their singular region and remove all sorts of artifacts from the primary skeleton is still another challenging topic.

Although the skeleton is an attractive representation of character, its application to image processing and pattern recognition is limited. Therefore, the improvement and amendment processing of the primary skeleton also play an important role in the implementation of the skeletonization.

In addition, most of the existing methods are applicable to only the clean binary images rather than the gray-level images containing much information of the character. Meanwhile, these methods are sensitive to the noise and affine transformations.

The main contributions of this thesis are as follows: A new symmetry analysis is proposed to directly extract the primary skeleton. This symmetry analysis is based on the desirable properties of the new wavelet function, which has been constructed in our previous work [11, 12]. The points of the primary skeleton can be detected straightforward by computing the modulus maxima of the wavelet transform, which is ease to implement.

The proposed thesis focuses on developing a simple and efficient amendment technique using wavelet-based interpolation compensation to extract the skeleton points in the singular region of the character strokes. The key of this interpolation compensation is to position the skeleton points in the junction of the singular region in terms of three types of characteristic points. To measure these characteristic points, multiscale-based technique for corner detection is also improved by using new wavelet function.

3. METHODOLOGY

3.1 Wavelet-Based Symmetry Analysis

Symmetry analysis of the regular region depends on the detection of a pair of contour curves in the regular region of the stroke (or a pair of edges). In this thesis, a wavelet-based edge detection, which is the development of the traditional one (such as canny's method), is applied to detect a pair of contour curves.

Wavelet-based edge detection is to extract the local maxima of the wavelet transform moduli. The wavelets imply the partial derivatives of a lowpass filter.

As far as the skeleton extraction is concerned, the desired wavelet not only locates the edge points by the local maxima of the wavelet transform moduli, but also adjust properly the position of local maxima around the exact edge points, in order to find easily the center of the regular region of a stroke by a symmetric pair of local maxima.

The proposed methodology, construct a wavelet, in which a pair of contour curves extracted by the corresponding wavelet transform is located outside the original edges of the shape, and the middle axis (skeleton) of the shape is the symmetrical axis of these contour curves. In addition, the

distance between these two contour curves depends strongly on the scale of the wavelet transform.

The purpose of smoothing is to remove the noise, not the edge points. For this reason, it should be localized, such that it will be small enough. This implies that the smoothed signal is almost the same as the original one, when the scale is very small.

3.2 Algorithm of Extraction of Primary Skeleton

The primary skeleton of the regular region is made by the sequence of all the mid-points of a symmetric pair of the modulus maxima of the wavelet transform. Theoretically, given a modulus maximum point of the wavelet transform, it can search another maximum point as its symmetric counterpart along its gradient direction such that the distance between them equals to the scale.

In the discrete domain, the gradient directions are quantized by eight gradient codes (GCs) according to the tangent of a gradient direction. Therefore, a plane can be divided into eight sectors. When falls into a sector, it will be quantified to a certain vector, which is represented by a central line of that sector. It indicates the direction of that gradient. For every modulus maxima point, even though the corresponding symmetric counterpart cannot be achieved or not exist along its gradient direction, one or more modulus maxima points may appear in a sector, which contains certain GC. Hence, from these points, we can select one as the symmetric counterpart, where the distance between this point and the original modulus maxima point approximates scale.

3.3 Smooth Processing

The smooth processing is done by examining the gradient code (GC) of its neighbor skeleton points along the normal direction of its gradient. Namely, if it has the same GC with its neighbor skeleton points, it is considered to be a skeleton point. Some lost skeleton points are complemented, and the amended skeletons are continuous and smooth and conform to human perceptions. They are helpful to further computing the skeletons of the singular region as well. This amendment process is a smooth processing, and the corresponding amended skeleton is a continuous segment, which is called the skeleton segment.

In terms of the smooth processing, two types of contour points of the character strokes can be obtained i.e., stable and unstable contour points, which will be used for the classification of the characteristic points.

Stable contour point:

If the contour point has a counterpart as modulus maxima symmetric pair of wavelet transforms or its corresponding skeleton point can be obtained by the smooth processing. The segment of the stable contour consists of consecutive stable contour points.

Unstable contour point:

The contour point does not have a counterpart as a modulus maxima symmetric pair of wavelet transform, or its corresponding skeleton point cannot be obtained by the smooth processing. The segment of the unstable contour consists of consecutive unstable contour points.

3.4 Interpolation Compensation

The singular regions of the character strokes can be categorized into five typical patterns, namely, T-pattern, Y-pattern, X-pattern, V-pattern, and K-pattern. In order to compute the skeleton points of the singular regions of the strokes, the system develop a wavelet-based interpolation compensation technique. The key of this technique is to position the junction skeleton points in the singular region by detecting three types of characteristic points: terminal point, divider point, and corner point.

The junction skeleton point can be estimated by two ways:

a) It is the centroid of the polygon, which is composed of the above three types of characteristic points.

b) The position of the junction skeleton point can be calculated by the least square method with the characteristic points.

The corner point of the planar curve is a common concept, it is not necessary to redefine. The start or end point of the skeleton segment obtained from smooth processing is defined as terminal point. The intersection point between the stable and the unstable contour segments is called divider point. The terminal points are the start and end points of the skeleton segment, and can be easily obtained from the skeleton segment after smooth processing. Along the gradient direction of the terminal

point, the divider points can be detected easily. Thus, the main task of the detection of the characteristic point aims at computing the corner point.

4. SYSTEM MODEL

The system model comprises of the following component

4.1 Image Preprocessing

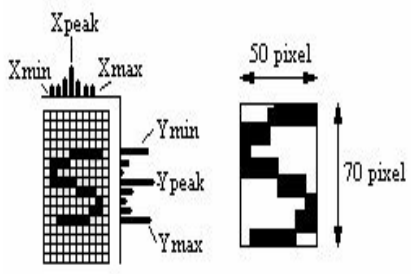
The image is first being converted to grayscale image follow by the threshing technique, which make the image become binary image. The binary image is then going through connectivity test in order to check for the maximum connected component, which is, the box of the form. After locating the box, the individual characters are then cropped into different sub images that are the raw data for the following feature extraction routine.

The size of the sub-images are not fixed since they are expose to noises which will affect the cropping process to be vary from one to another. This will causing the input of the network become not standard and hence, prohibit the data from feeding through the network. To solve this problem, the sub-images have been resize to 50 by 70 and then by finding the average value in each 10 by 10 blocks, the image can be down to 5 by 7 matrices, with fuzzy value, and become 35 inputs for the network. However, before resize the sub-images, another

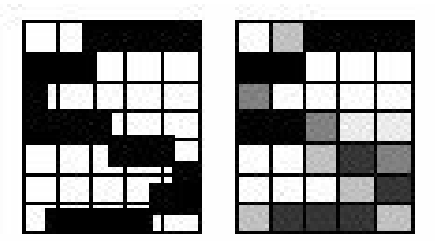
process must be gone through to eliminate the white space in the boxes.

4.2 Feature Extraction

The sub-images have to be cropped sharp to the border of the character in order to standardize the sub-images. The image standardization is done by finding the maximum row and column with 1s and with the peak point, increase and decrease the counter until meeting the white space, or the line with all 0s. This technique is shown in figure below where a character "S" is being cropped and resize.



The image pre-processing is then followed by the image resize again to meet the network input requirement, 5 by 7 matrices, where the value of 1 will be assign to all pixel where all 10 by 10 box are filled with 1s, as shown below:



Finally, the 5 by 7 matrices are concatenated into a stream so that it can be feed into network 35 input neurons. The input of the network is actually the negative image of the figure, where the input range is 0 to 1, with 0 equal to black and 1 indicate white, while the value in between show the intensity of the relevant pixel.

4.3 Extraction of primary skeleton

To extract the primary skeleton in the regular region of the character stroke, the key technique is to construct a local symmetry of the contour of the regular region, and to compute the symmetric center points, which produce the skeleton. Hence, the methodology first focus on developing a new symmetric analysis based on the modulus maxima of the wavelet transform.

Virtually, the process of computing the wavelet transform and detecting the modulus maxima is equivalent to contour extraction in the traditional algorithms. As long as the modulus maxima are computed, the skeleton can be computed straightforward as the byproducts of the process of the foregoing contour detection. Hence, the consumed time in the skeleton extraction step may be ignored compared with the contour detection. Therefore, the implementation of the proposed algorithm is expected to be easy, direct and fast.

The algorithm is robust against the noise and distracting background. The proposed algorithm is applicable to both the binary and gray level images (most existing algorithms fail to process the gray level image). As we have known, the noise always distributes randomly. From the statistical point of view, the average value of the noise is nearly a constant in a certain area. In general, it can suppose that the value of this constant is zero, and take a weighted mean to the signal in this area. This action can be regarded as a low-pass filtering. In this way, the noise will be eliminated considerably.

4.4 Amendment processing of singular region of stroke

The proposed amendment processing consists of two major operations i.e., smooth processing and interpolation compensation. The primary skeleton obtained deals with only the regular regions of a stroke. This primary skeleton usually seems to be rough with poor visual appearance since the skeleton loci are not continuous and not smooth.

This is because that the symmetric counterparts of some contour points of the segments of the regular region cannot be found due to some factors, such as the image digitization, approximation, noise, computation error, and so on. Therefore, some skeleton points are lost from the primary skeleton, which may produce some gaps. This gap problem can be easily amended by using the smooth technique.

A key issue of our amendment processing is to compute the skeletons of the singular region of the stroke. Some skeleton points in the singular region of the character stroke cannot be extracted using the primary skeleton algorithm. A wavelet-based amendment processing called interpolation compensation technique will be developed to realize this task.

5. IMPLEMENTATION AND EXPERIMENTAL EVALUATION

5.1 Implementation

First obtain all characteristic points in the singular region of the strokes, and then related characteristic points will be connected to produce a polygon. The junction skeleton point in the singular region can be determined by the centroid of such a polygon or by least square method.

In practice, a polygon is generated as follows: For each terminal point in the skeleton segment, along its gradient (positive and negative) direction, one of two divider points should be found. Along its corresponding unstable contour segment, another divider point (or corner point) can be found. All these characteristic points related to the terminal point are recorded and form a subgroup.

When two subgroups share some characteristic points, two subgroups are merged into a bigger subgroup. The same search continues until a new characteristic point is included. Finally, all these subgroup generate a union group, which includes all characteristic points in the singular region, and they can be connected to construct a polygon.

5.1.1 Algorithm for extracting the primary skeletons

a) Estimate the maximum width of the strokes of character image by applying elastic meshing method.

b) Select a suitable scale according to the above maximum width and perform the wavelet transform to the character image.

c) Compute the modulus of the wavelet transform and the gradient direction.

d) Threshold the modulus image to remove the noise (if the noise exists).

e) Compute the local modulus maxima.

f) Find the symmetric pair of modulus maxima of a segment of the regular region of the stroke, and thereafter determine the primary skeleton points in terms of the midpoints of symmetric contour pair of the segment.

5.1.2 Wavelet-based corner detection

a) To obtain the contour of the character strokes by computing the modulus maxima of the wavelet transform.

b) Compute the orientation function of every contour.

c) Apply the wavelet transform to the orientation at two scales. The moduli , are stored in the one dimension (1-D) array.

d) The contour points, at which moduli at two consecutive scales simultaneously reach the maxima, are detected as the candidates of the corner points.

5.2 Experimentation

Apply the proposed method to English characters, which are represented by both the binary images and gray images. In the experiments, first the modulus maxima of the wavelet transform of the original image will be obtained, and thereafter, the primary skeleton will be produced by applying the proposed amendment algorithm, finally, the final skeleton can be generated by the proposed amendment processing.

Skeletonization problems often occur on coarsely sampled input. To examine the effectiveness, the proposed method is tested on the coarsely sampled images. The raw modulus maxima of the wavelet transform with scale should be displayed. The primary results and the final skeletons obtained from the proposed method are applicable for images sampled at low resolution.

6. CONCLUSION

The main step of the algorithm is built on the polygonal line algorithm which approximates principal curves of data sets by piecewise linear curves. The algorithm also contains two operations specific to the task of skeletonization, an initialization method to capture the approximate topology of the character, and a collection of restructuring operations to improve the structural quality of the skeleton produced by the initialization method.

Test results indicate that the proposed algorithm can be used to substantially improve the pixel wise skeleton obtained by traditional thinning methods. Results on continuous handwriting demonstrate that the skeleton produced by the algorithm can also be used for representing handwritten text efficiently and with a high accuracy.

Extraction of the primary skeleton of the regular stroke depends on the new symmetric analysis of the modulus maxima of the wavelet transform. The implementation of the new symmetry analysis benefits from the desirable characteristics of the new wavelet function. Under the wavelet transform, a pair of contour curves (local maxima of wavelet transform moduli) of the regular stroke is located outside the edges of the stroke and is symmetric with respect to the central line of the regular stroke.

Meanwhile, the distance between the symmetric points equals to scale of the wavelet transform. The amendment process includes the smooth processing and interpolation compensation. The smooth processing is based on the congruence of the gradient code and is applied to amend the lost skeleton points in the regular regions of the character strokes.

The interpolation compensation is used to compute the skeletons in the singular regions of the strokes. Its essential operation is to position the junction point of the skeleton from three types of the characteristic points and then connect the junction skeleton point with the terminal points to finalize the skeleton.

REFERENCES

[1] L. Lam, S. W. Lee, and C. Y. Suen, "Thinning methodologies comprehensive survey," IEEE Trans. Pattern Anal. Mach. Intell., vol. 14, no. 9, pp. 869–885, Sep. 1992.

[2] R. W. Smith, "Computer processing of line images: A survey," Pattern Recognit., vol. 20, pp. 7–15, 1987.

[3] E. S. Deutsch, "Preprocessing for character recognition," in Proc. IEEE NPL Conf. Pattern Recognition, 1968, pp. 179–190.

[4] B. Kégl and A. Krzyÿzak, “Piecewise linear skeletonization using principal curves,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 1, pp. 59–74, Jan. 2002.

[5] H. Blum, “A transformation for extracting new desxriptors of shape,” in Models for the Perception of Speech and Visual Form, W. Wathen Dunn, Ed. Cambridge, MA: MIT Press, 1967, pp. 362–380.

[6] M. Leyton, “A process-grammar for shape,” Artif. Intell., vol. 34, pp. 213–247, 1988.

[7] J. C. Simon, “A complemental approach to feature detection,” in from Pixels to Features, J. C. Simon, Ed. Amsterdam, The Netherlands: North-Holland, 1989, pp. 229–236.

[8] H. Blum, “Biological shape and visual science (Part 1),” Biology, 1973. [9] C. Wang, D. J. Cannon, R. T. Kumara, and G. Lu, “A skeleton and neural network-based approach for identifying cosmetic surface flaws,” IEEE Trans. Neural Netw., vol. 6, no. 5, pp. 1201–1211, Sep. 1995.

[10] J. J. Zou and H. Yan, “Skeletonization of ribbon-like shapes based on regularity and singularity analyses,” IEEE Trans. Syst. Man. Cybern. B, Cybern., vol. 31, no. 3, pp. 401–407, Jun. 2001.

[11] Y. Y. Tang and X. G.You, “Skeletonization of ribbon-like shapes based on a new wavelet function,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no. 9, pp. 1118–1133, Sep. 2003.

[12] X. You, Y. Y. Tang, and L. Sun, “Skeletonization of ribbon-like shapes with new wavelet function,” in Proc. 1st Int. Conf. Machine Learning and Cybernetics, 2002, pp. 1869–1874.

EXTENDED VISUAL CRYPTOGRAPHY USING COMPLIMENTARY SHARE PARTICIPANT VERIFICATION

J. Nirmala Devi
II M.E CSE
Department of Computer Science
and Engineering
K.S.Rangasamy College of Technology,
KSR Kalvi Nagar,
Thiruchengode-637 209
Tamil Nadu, India.
91-9843918686
yasavi_sudhan@yahoo.co.in

ABSTRACT

The ability to share all kinds of information and resources is fundamental in today's global environment. However, at the same time a whole variety of security systems using encryption methods have also been developed to prevent information from being accessed or used by unauthorized people. Visual cryptography is a very powerful method for secret sharing and encrypting information, especially an image. However, the resulting encrypted image has low contrast and signal-to-noise ratio (SNR) because of the sub pixels generated during the encryption process. Extended Visual Cryptography (EVCS) is a type of cryptography which encodes a number of images in the way that when the images on transparencies are stacked together, the hidden message appears without a trace of original images. The decryption is done directly by the human visual system with no special cryptographic calculations. The proposed thesis develop a system which takes three pictures as an input and generates two images which correspond to two of the three input pictures. The third picture is reconstructed by printing the two output images onto transparencies and stacking them together. While the previous researches basically handle only binary images, this thesis establishes the extended visual cryptography scheme suitable for natural images. Visual cryptography suffers from the deterioration of the image quality. The proposed system describes the method to improve the quality of the output images. The trade-off between the image quality and the security are discussed and assessed by observing the actual results of this method. Furthermore, the optimization of the image quality is discussed.

Categories and Subject Descriptors

I.4.6 Segmentation [Pixel classification], I.4.m Miscellaneous.

General Terms

Algorithms, Documentation, Security, Verification.

Keywords

Visual cryptography, Extended Visual Cryptography, transparency.

1. INTRODUCTION

Visual cryptography (VC) is a kind of cryptography that can be decoded directly by the human visual system without any special calculation for decryption. The VC system takes three pictures as an input and generates two images which correspond to two of the three input pictures.

The third picture is reconstructed by printing the two output images onto transparencies and stacking them together. This type of visual cryptography, which reconstructs the image by stacking some meaningful images together, is especially called Extended Visual Cryptography. In this thesis, the pictures shown on the output images are called sheets and the resulting image reconstructed by stacking the two sheets together is called the target.

Visual Cryptography encrypts a secret image into several shares but requires neither computer nor calculations to

decrypt the secret image. Instead, the secret image is reconstructed visually, by overlaying the encrypted shares the secret image becomes clearly visible. A visual cryptography scheme [VCS] is a visual secret sharing scheme such that stacking any or more shares reveals the secret image, but stacking fewer than shares reveals not any information about the secret image. VC has been studied intensively since the pioneer work of Noar and Shamir [9].

Most of the previous research work on VC focused on improving two parameters: pixel expansion and contrast [4,5].In these cases, all participants who hold shares are assumed to be semi-honest, that is, they will not present false or fake shares during the phase of recovering the secret image. Thus, the image shown on the stacking of shares is considered as the real secret image. Nevertheless, cryptography is supposed to guarantee security even under the attack of malicious adversaries who may deviate from the scheme in any way. It is possible to cheat [6, 7] in VC, though it seems hard to imagine.

For cheating, a cheater presents some fake shares such that the stacking of fake and genuine shares together reveals a fake image. With the property of unconditional security, VC is suitable for sending highly classified orders to a secret agent when computing devices may not be available.

The secret agent carried some shares, each with a pre-determined order, when departing to the hostile country. When the head quarter decides to execute a specific order, it can simply send another share to the agent so that the agent can recover what the order is. It can be seen that it would be terrible if the dispatched share cannot be verified due to a cheater's attack.

A VCS would be helpful if the shares are meaningful or identifiable to every participant. A VCS with this extended characteristic is called extended VCS (EVCS) [2], [9]. An EVCS is like a -VCS except that each share displays a meaningful image, which will be called share image hereafter. Different shares may have different share images. At first glance, it seems very difficult to cheat in EVCS because the cheater does not know the share images that appear on the genuine shares and thus, has no information about the distributions of black and white pixels of the share images. This information is crucial for cheating in VC.

The proposed thesis, discuss the cheating problem in VC and EVC. It present cheating methods i.e., malicious participant and apply them on EVC schemes. The attacks are to reveal fake images to cheat honest participants. The attacks are more like the man-in-the-middle attack in cryptography.

The proposal develops a generic method that converts a VCS to another VCS that has the property of cheating prevention. The overhead of the conversion is optimal. The digressions in contrast of the converted VCS are almost optimal.

2. LITERATURE REVIEW

Naor and Shamir [9] proposed a -VCS. Many improvements and extensions follows [1, 3]. For example, Ateniese et al. [1] proposed an elegant VCS for general access structures based on the cumulative array method. Tzeng and Hu [10] proposed a new definition for VC, in which the secret image can be either darker or lighter than the background.

Naor and Pinkas [8] showed some methods of authentication and identification for VC. Their scenario focuses on authentication and identification between two participants. Yang and Laih [11] proposed two cheat-preventing methods. Their first method needs an on-line TA (Trusted Authority) to verify the shares of participants. Their second method is a transformation from a VCS (but not a -VCS) to a cheat-preventing VCS on which the stacking of two shares reveals the verification image. The method needs to add extra sub pixels for each pixel in the secret image.

Horng et al. [7] proposed a cheating method against some VC schemes. In their cheating method, the cheater needs to know the exact distribution of black and white sub pixels of the shares of honest participants. Based on this characteristic, they proposed a cheat-preventing method to prevent the cheater from obtaining the distribution. However, we show that the knowledge of the distribution is not a necessary condition for a successful cheat. They also proposed another cheat-preventing method in which the stacking of the genuine share and verification share reveals the verification image in some small region.

3. METHODOLOGY

3.1 Visual Secret Sharing Scheme

The basic model of the visual cryptography consists of a several number of transparency sheets. On each transparency a cipher text is printed which is indistinguishable from random noise. The hidden message is reconstructed by stacking a certain number of the transparencies and viewing them. The system can be used by anyone without any knowledge of cryptography and without performing any cryptographic computations.

The system has developed the Visual Secret Sharing Scheme (VSSS) to implement this model. In k out of n VSSS (which is also called (k, n) scheme), a binary image (picture or text) is transformed into n sheets of transparencies of random images. The original image becomes visible when any k sheets of the n transparencies are put together, but any combination of less than k sheets cannot reveal the original binary image.

In the scheme, one pixel of the original image is reproduced by m sub pixels on the sheets. The pixel is considered "on" (transparent) if the number of transparent sub pixels is more than a constant threshold, and "off" if the transparent sub pixels is less than a constant lower threshold, when the sheets are stacked together. The contrast is the difference between the on and off threshold number of transparent pixels.

It has extended the (k, n) VSSS to general access structures where senders can specify all qualified and forbidden subsets of n participants. It considered the problem of sharing more than one secret image among a set of participants, and proposed a method to reconstruct different images with different combinations of sheets.

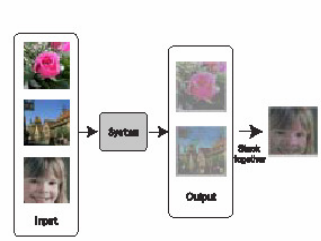


Figure 1: System Model

3.2 Extended Visual Cryptography

The extension of the VCS is proposed which conceals the every existence of the secret message. That is, each sheet carries some meaningful images rather than random dots. They referred to the (2, 2) example with the number of sub pixels m = 4. It has formalized this framework as the Extended Visual Cryptography and developed a scheme for general access structures. It discuss the trade-off between the contrast of the each images on the sheets and that of the resulting image when stacked together in (k, k) cases.

The proposed scheme, have discussed the visual cryptography for grayscale and color images. It is extended to grayscale images i.e., to represent the gray levels of the hidden image by controlling the way how the opaque sub pixels of the sheets are stacked together.

The VCS can deal with color images as well. It discussed the visual cryptography scheme which reconstructs a message with two colors, by arranging the colored or transparent sub pixels. It assigns a color to a sub pixel at a certain position, which means that displaying m colors uses m-1 sub pixels. The resulting pixels contain one colored sub pixel and the rest of the sub pixels are black. Therefore the more colors are used, the worse the contrast of the images becomes significantly.

4. SYSTEM MODEL AND EXPERIMENTAL EVALUATION

The system model consists of two phase namely generating cheating sequence in EVCS and developing a more generalized Cheating prevention scheme with the high quality output image.

4.1 Cheating in EVCS

There are two types of cheaters in our scenario. One is a malicious participant who is also a legitimate participant, and the other is a malicious outsider. A cheating process against a VCS consists of the following two phases:

1) Fake share construction phase - the cheater generates the fake shares,

2) Image reconstruction phase - the fake image appears on the stacking of genuine shares and fake shares.

In order to cheat successfully, honest participants who present their shares for recovering the secret image should not be able to distinguish fake shares from genuine shares. A reconstructed image is perfect black if the sub pixels associated to a black pixel of the secret image are all black. Most proposed VC schemes have the property of perfect blackness. For example, the reconstructed secret images are all perfectly black.

It only considers cheating the participants who together do not constitute a qualified set. Since all participants together in a qualified set can recover the real secret image in perfect blackness already, it is not possible to cheat them.

Since the cheater is a malicious participant, he uses his genuine share as a template to construct a set of fake shares which are indistinguishable from its genuine share. The stacking of these fake shares and S1 reveals the fake image of perfect blackness.

In the definition of VC, it only requires the contrast be nonzero. Nevertheless, we observe that if the contrast is too small, it is hard to “see” the image. Based upon this observation, the idea of cheating method is to use the fake shares to reduce the contrast between the share images and the background. Simultaneously, the fake image in the stacking of fake shares has enough contrast against the background since the fake image is recovered in perfect blackness.

4.2 Cheating Prevention on EVCS

There are two types of cheat-preventing methods. The first type is to have a trusted authority (TA) to verify the shares of participants. The second type is to have each participant to verify the shares of other participants.

The first cheat-preventing method needs a Trusted Authority to hold the special verification share for detecting fake shares. The verification images should be confidential and spread over the whole region of a share.

It uses a VCS to implement a VCS cheat-preventing scheme. The scheme needs no on-line TA for verifying shares.

The scheme generates shares by the VCS for some integer, but distributes only shares to the participants. The rest of shares are destroyed. They reason that since the cheater does not know the exact basis matrices even with all shares, the cheater cannot succeed. However the cheating method, do not need to use the basis matrices. The second cheat-preventing method is a transformation of a VCS to another cheat-preventing VCS. The stacking of any two shares reveals the verification image. This is how share verification is done.

The proposed generic cheat prevention method work on the following aspects:

a) It does not rely on the help of an on-line TA. Since VC emphasizes on easy decryption with human eyes only, it should not have a TA to verify validity of shares.

b) Each participant verifies the shares of other participants. This is somewhat necessary because each participant is a potential cheater.

c) The verification image of each participant is different and confidential. It spreads over the whole region of the share. It is shown that this is necessary for avoiding the described attacks.

d) The contrast of the secret image in the stacking of shares is not reduced significantly in order to keep the quality of VC.

Each participant holds a verification share. The cheat-preventing scheme needs no help from an on-line TA. The verification image for each participant is different and known to the participant only. The transformation is quite efficient and almost optimal as it adds only two sub pixels for each pixel of the original image.

The transformation generates two shares for each participant. One is the secret share and the other is the verification share. The transformed VCS use the basis matrices to generate shares for the participants as usual. Then, for each participant, it generates a verification share for a chosen verification image. The verification image is encoded into the first two sub pixels. If participant wants to verify the share of participant, he checks whether it shows his verification image.

4.3 The Contrast Enhancement

The proposed scheme aims at natural images, eg, grayscale images, with higher image quality. It analyzes and evaluates the contrast enhancement as a way of improving the quality. First the proposed method applies affine transformation to the each pixel's intensity (transparency) in order to reduce the contrast of the input images, since though contrast can be enhanced; it is difficult to encrypt the input images themselves.

For simplicity, suppose that the contrast of the sheets and that of the target are the same fixed value K. It is obvious that 0 is the most appropriate value for the lower bound of the target because the target must be darker than both sheets. Let L denote the lower bound of the sheet dynamic range.

When enhancing the contrast, it is necessary to consider the condition for the encryption. It is sufficient for encryption with any three arbitrary images. In fact, the encryption can be performed if, in each triplet (pixels of two sheets and the target).

For binary images, the conflicts are fatal to the encryption. However, grayscale images can tolerate those conflicts by adjusting gray levels of the conflicting triplets. The proposing method performs both half-toning and encryption process simultaneously to enable this adjustment with natural results. As the contrast enhances, the conflicts with the conditions are adjusted and the resulting errors are diffused to the nearby pixels.

5. CONCLUSION

The thesis developed a generic cheating prevention scheme for extended visual cryptography with high quality output images. The cheating model involves malicious participant and malicious outsider. The system proposed an efficient transformation of VCS for cheating prevention. The transformation incurs minimum overhead on contrast and pixel expansion. It only added two sub pixels for each pixel in the image and the contrast is reduced only slightly.

The proposal of extended visual cryptography scheme for the binary images showed a method to improve the image quality of the output by enhancing the image contrast beyond the constraints given by the previous studies. The method enables the contrast enhancement by extending the concept of error and by performing half toning and encryption simultaneously.

The trade-off between the image quality and the security are assessed by observing the actual results of this method. The optimization of the image quality at a given contrast is also analyzed. The validity of the assumption and the effect of image quality improvement are also verified with the experiments.

REFERENCES

[1] G. Ateniese, C. Blundo, A. De Santis, and D. R. Stinson, "Visual cryptography for general access structures," Inf. Comput., vol. 129, no. 2, pp. 86–106, 1996.

[2] "Extended capabilities for visual Cryptography," Theoret. Comput. Sci., vol. 250, no. 1–2, pp. 143–161, 2001.

[3] I. Biehl and S.Wetzel, "Traceable visual Cryptography," in Proc. 1st Int. Conf. Information Communication Security, 1997, vol. 1334, LNCS, pp. 61–71.

[4] C. Blundo, A. De Santis, and D. R. Stinson, "On the contrast in visual cryptography schemes," J. Cryptol., vol. 12, no. 4, pp. 261–289, 1999.

[5] C. Blundo, P. D'Arco, A. De Santis, and D. R.Stinson, "Contrast optimal threshold visual Cryptography schemes," SIAM J. Discrete Math., vol. 16, no. 2, pp. 224–261, 2003.

[6] H.Yan, Z. Gan, and K. Chen, "A cheater detectable visual cryptography scheme," (in Chinese) J. Shanghai Jiaotong Univ., vol. 38, no. 1, 2004.

[7] G.-B. Horng, T.-G. Chen, and D.-S. Tsai, "Cheating in visual cryptography," Designs, Codes, Cryptog., vol. 38, no. 2, pp. 219–236, 2006.

[8] M. Naor and B. Pinkas, "Visual authentication and identification," in Proc. Advances in Cryptology, 1997, vol. 1294, LNCS, pp. 322–336.

[9] M. Naor and A. Shamir, "Visual Cryptography," in Proc. Advances in Cryptology, 1994, vol. 950, LNCS, pp. 1–12.

[10] W.-G. Tzeng and C.-M. Hu, “A new approach for visual cryptography,” Designs, Codes, Cryptog., vol. 27, no. 3, pp. 207–227, 2002.

[11] C.-N. Yang and C.-S. Laih, “Some new types of visual secret sharing schemes,” in Proc. Nat. Computer Symp., 1999, vol. 3, pp. 260–268.

Design & Development of Fuzzy Logic Based Embedded Controller by using HDL with LCD interfacing

Shakti Kumar , Amrit Kaur,
Institute for Science & Technology Klawad , (Haryana) INDIA
Rayat & Bahara Institue of Engineering & Biotechnology Saharaun, Kharar

Abstract: **This paper presents a design & development of fuzzy logic based systems using Hardware Descriptive Languages i.e VHDL & Verilog. Design automation gets the design requirements for the system to be designed from the user. Then generates the VHDL & Verilog code for the complete system. This code can then be synthesized and implemented onto a CPLD/FPGA using XILINX or any other software.It slashes the design time considerably and brings the design into the market within any permissible time span.**

I INTRODUCTION

Embedded controller is first design and then developed by using VHDL and Verilog . Fuzzy systems are a class of knowledge based systems. Owing to their advantages these systems have received a considerable attention in the recent past. Recent market trends can ill afford larger design cycle times. " Quick to market" is fundamental rule for the success of products The embedded controller is used to control the charging of the battery by using Fuzzy Controller. Design Automation provides the design of a rapid Nickel Cadmium battery charger. Rapid charging needs high charging current of the order of 8C where C represents the capacity of the battery. The battery charger needs an intelligent strategy as the charging has to be controlled to avoid damage to the battery due to rise in temperature caused by high charging currents. The temperature control assumes significance as the rise in temperature is exponential after initial charge. The boom in the electronics industry has revolutionized the field of battery management, which includes regulating the charging, protection, and monitoring of the battery because the batteries limit the performance of the electronic devices. Keeping view all the above points, the need to accentuate research in the field of battery management is clear enough.

II EMBEDDED SYSTEM

The present day life is completely overwhelmed by the magnificence of the embedded systems and their applications. Starting right from the kitchen of a common housewife to the astronauts' space carriers and a researchers' personal computer or laptop, all facets of life have been touched and then captured by the embedded systems. A designer is always constrained to complete his or her design in the minimal possible time. It is well said that he who launches the product first wins the market. Hence, it becomes very important for a designer to reduce the design time to the bare minimum, allowing extensive testing of the designed systems in order to reduce the chances of failure if any. In order to reduce the design time it becomes essential for a designer to search for or design some automated design tool for the embedded system to be designed. The present work is a step in this direction.

III HARDWARE DESCRIPTIVE LANGUAGE

A hardware description language (HDL) is a software programming language used to model the intended operation of a piece of hardware .Design idea to the EDA tool in such a way that the tool can understand that and generate a highly optimized design implementation. This transfer of idea from human beings to machines is called design entry. There are two methods of design entry: graphical and textual. Graphical method includes schematic entry and state diagram entry while textual method implies use of HDLs. Graphical method works fairly well up to a certain level of circuit complexity but in larger circuits they become cumbersome and time consuming[1].In this paper we are using textual method.

Advantages of HDL

- Compact description
- Easy to edit

- Highly portable
- Rapid prototyping of design
- Availability of extensive vendor libraries.
- Increasing capability of synthesis tools

Limitations of HDL

- No support for analog behaviour
- Need to learn coding styles that ensures synthesizable results
- Design Methodology

IV FUZZY LOGIC

The use of Fuzzy logic has received increased attention in the recent past because of its ability to deal with highly complex and nonlinear systems. Also it could reduce the need for complex mathematical model that may be very difficult to come by for complex systems. Fuzzy logic is a powerful problem-solving methodology with wide range applications in industrial control, consumer electronics, management, medicine, expert systems[6,7].

A. Necesssity of fuzzy logic

Fuzzy logic deals with uncertainty in engineering by attaching degrees of certainty to the answer to a logical question[1].

1) Better way of dealing with:
   (i) ill-defined systems
   (ii) Imprecise knowledge
   (iii) Incomplete information
   (iv) Complex & non-linear systems
2) No need for rigorous mathematical treatment.
3) Better way of dealing with linguistic expressions.
4) Can be blended with conventional control techniques.
5) Reduction of development & maintenance times.
6) Based on natural language.

B .System Design

This section discusses the design and implementation of various modules of a fuzzy processor on the microcontrollers. To design a fuzzy logic based embedded controller, the following steps are usually followed [11, 13]:

1. Identify the inputs and outputs variables
2. Partitioning of input and output space i.e., assigning membership functions to the variables(fuzzy subsets; fuzzification
3. Building a rule base
4 .Reducing the rule base if possible.
5. Choosing an inference strategy.
6 .Selecting defuzzification scheme.)

C. Modules of Fuzzy Logic

A Fuzzy system can be represented with the help of a block diagram as shown in figure Any Fuzzy system consists of four major modules of the system [13].

i). Fuzzification Module
ii) Inference engine
iii) Knowledge base
iv) Defuzzification module

The fuzzification module transforms the crisp input(s) into fuzzy values. These values are then processed in a fuzzy domain by inference engine based on the base (rule base and procedural knowledge) supplied by the domain experts[4]. Finally the processed output is transformed from fuzzy domain to crisp domain by defuzzification module.

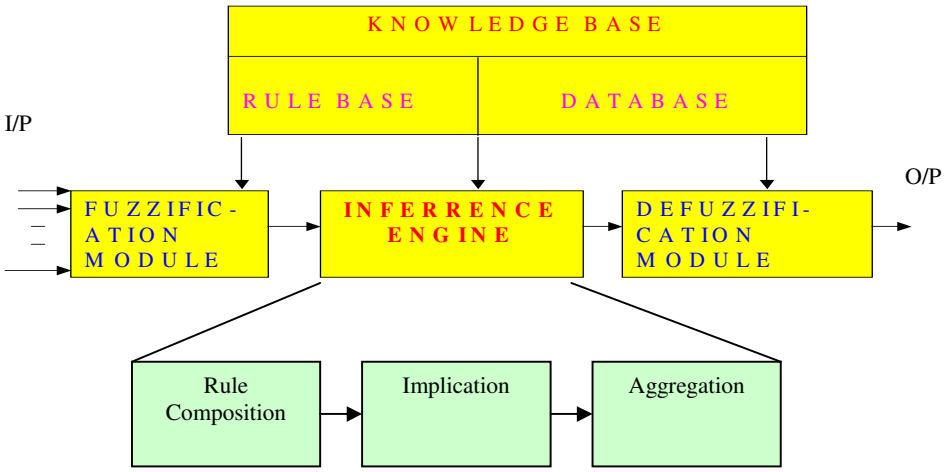


Fig 1. Block Diagram of Fuzzy system

i) Fuzzification Module

Fuzzification is related to the vagueness and imprecision in a natural language. It is a subjective valuation, which transforms a measurement into a valuation of an objective value, and hence it could be defined as a mapping from an observed input space to fuzzy sets in certain input universes of discourse. In fuzzy control applications the observed data are usually crisp. Since the data manipulation in an FLC is based on fuzzy set theory, fuzzification is necessary in an earlier age. The fuzzification module (FM) performs the following functions:

1) Measures the value of input variables.
2) Receives the crisp inputs.The input variables are scale mapped that transfers the range of values of input variables into corresponding universe of discourse.
3) Transforms the input crisp variables to fuzzy sets with corresponding membership grade.

a) Implementation of a Fuzzifier

This design and development tool uses four types of fuzzy sets namely Z-shaped, S-shaped, triangular type and trapezoidal or pi-type. The user can select any combination of these membership functions for any input as per system requirement. The universe of discourse for any input can be divided into maximum of 5 fuzzy sets.

In general ***Z-function*** can mathematically be expressed as:

$$\mu_Z(x) = \begin{cases} 1 & ; x \le a \\ (b-x)/(b-a) & ; a<x<b \\ 0 & ; b \le x \end{cases}$$

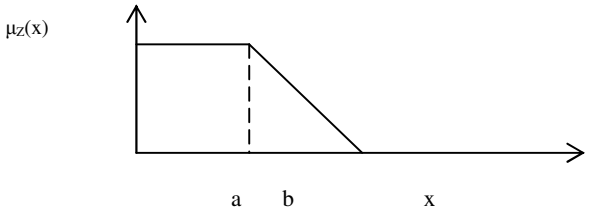


Fig 2. Z-membership function

where ***a*** and ***b*** are points of discontinuities as shown in figure 2. For any fuzzy set A, $\mu_A(x) \in [0,1]$ ie. the membership function maps the elements of fuzzy set A to real numbers from 0 to 1. As the system is implemented using microcontroller; where real valued operations are not possible, the whole range is scaled by FF hexadecimal, thereby giving formula for implementation of Z function as under. For ***S-shaped*** (fig 3) function we can write:

$$\mu_Z(x) = \begin{cases} FFh & ; x \le a \\ [(b-x) * FFh]/(b-a) & ; a<x<b \\ 0 & ; b \le x \end{cases}$$

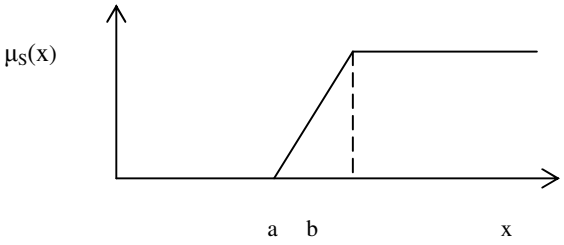


Fig.3 Sigma membership function

Similarly other membership functions are implemented using following formulae.

$$\mu_S(x) = \begin{cases} 0 & ;x \le a \\ [(x-a)*FFh]/(b-a) & ;a < x < b \\ FFh & ;b \le x \end{cases}$$

Triangular (fig 4) function can be e

$$\mu_T(x) = \begin{cases} 0 & ;x \le a \\ [(x-a)*FFh]/(b-a) & ;a<x<b \\ [(c-x)*FFh]/(c-b) & ;b<x \le c \\ 0 & ;c \le x \end{cases}$$

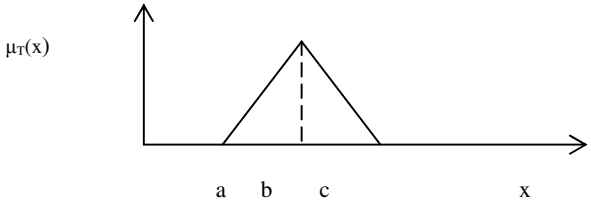


Fig 4 Triangular membership function

Similarly ***Pi-Type*** membership function (fig 5) can be expressed mathematically as follows:

$$\mu_P(x) = \begin{cases} 0 & ;x \le a \\ [(x-a)FFh]/(b-a) & ;a<x<b \\ FFh & ;b \le x<c \\ (d-x)FFh]/(d-c) & ;c \le x<d \\ 0 & ;d \le x \end{cases}$$

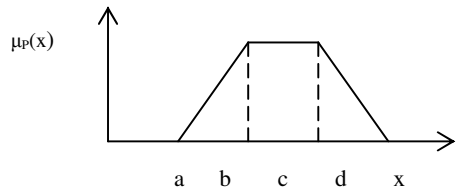


Fig5 Pi-type membership function

ii) INFERENCE ENGINE

Input	Membership grades from the previous stage i.e. from rule composition module.
Operation involved	1) MIN or MAX as t-norm or s-norm 2) Multiplication for Larsen's implication 3) Summation for aggregation

Fuzzy control rules are characterized by a collection of ***if-then*** rules in which the antecedents and consequents involve linguistic variables. The first stage of inference engine is rule composition. In rule composition some pre-conditions are specified in "***If***" part of rule. For rule composition AND or OR operators can be used. If OR operator is used then the composed value of rule is given as follows[5]:

$$\omega_k = \max \{\mu_i(x1), \mu_j(x2), \ldots\ldots\ldots\ldots \mu_m(xn)\}$$

where $\mu_m(xn)$ is the membership grade of n^{th} input in its m^{th} membership function. Similarly, for AND operator composed value of rule is given by:

$$\omega_k = \min \{\mu_i(x1), \mu_j(x2), \ldots\ldots\ldots\ldots \mu_m(xn)\}$$

Fuzzy rule base can be designed either by a domain expert or from the data. Three methods have been used for rule base generation. A classical method can be Wang and Mendel algorithm. GA's and PSO have also been used for rule base generation.

iii) Knowledge Base

Knowledge base consists of rule base and data base.

Data base provides the necessary data for linguistic variables and rule base generates the rules[4].

iv) Defuzzification

The last step in the design of a fuzzy system is the defuzzification. The defuzzification module transforms the output of the inference engine, which is a fuzzy set, into crisp data. Defuzzification can be treated as a generalized process.

The defuzzifier is implemented using the following weighted averaging formula.

$$\text{Defuzzified output} = \sum_{k=1}^{N} \omega_k * c_k \; / \; \sum_{k=1}^{N} \omega_k$$

MIN and MAX operators are also available to the designer as defuzzification modules.

many other methods as given below have been found in the literature. [18]

1. Weighted Average Method
2. Height defuzzification
3. Centroid /Centre of Gravity/Centre of Area Method
4. Centre of Sums
5. Centre of Largest Area
6. Max-Membership based Method
7. Middle (Mean) of Maxima
8. First (Last) of Maxima

.

The Developed verilog and VHDL code can presently be synthesized and simulated only using XILINX software as the code uses the DIVIDER component from XILINX component library .The Verilog implementation is limited to 10 inputs.

V. LIGHT CRYSTAL DIODE

LCD modules are cheap and easy to interface using a microcontroller or FPGA[2]. Here's a 1 line x 16 characters module.To control an LCD module, you need

11 I/O pins to drive an 8-bits data bus and 3 control signals(see Table 1). The 3 control signals are: E : enable, or "LCD-select". Active high. R/W: read/write. 0 to write, 1 to read.RS : register select, 0 for command bytes, 1 for data bytes. Most of the LCD modules are based on the HD44780 chip or compatible.LCD receives data from the PC serial port, de-serializes it, and send it to the LCD module. The de-serializer is the same module from the serial interface project

Table 1

Pin No	Symbol	Details
1	GND	Ground
2	Vcc	Supply Voltage +5V
3	Vo	Contrast adjustment
4	RS	0->Control input, 1-> Data input
5	R/W	Read/ Write
6	E	Enable
7 to14	D0 to D7	Data
15	VB1	Backlight +5V
16	VB0	Backlight ground

VI SIMULATED RESULT AND OUTPUT WAVEFORMS

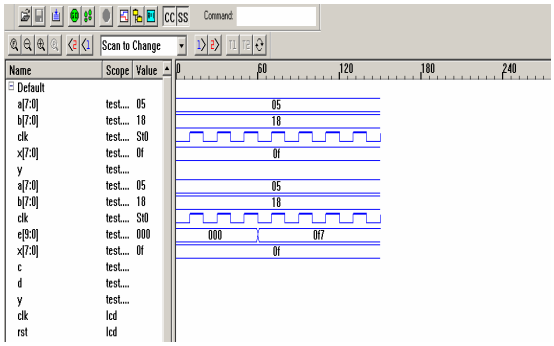


Fig.6 Simulated Waveform for Zed membership function

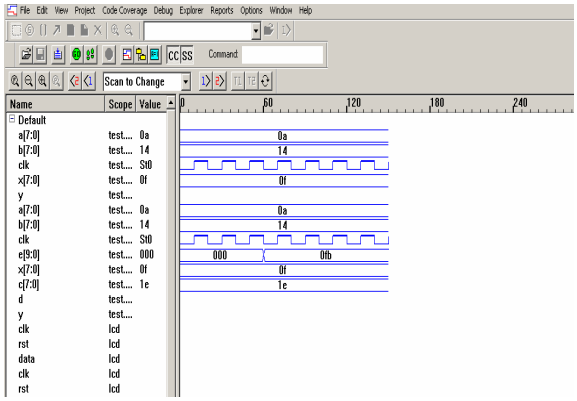


Fig7. Simulated Waveform for triangular membership function

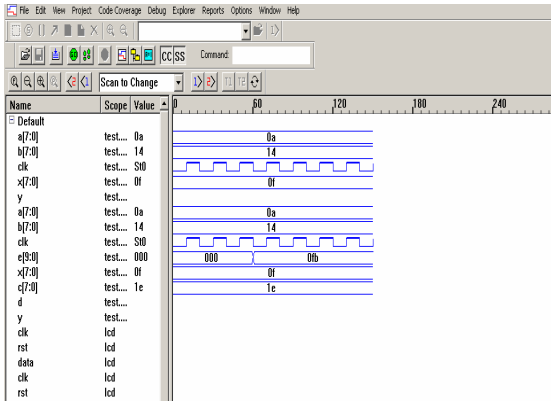


Fig8 Simulated Waveform for Triangular function

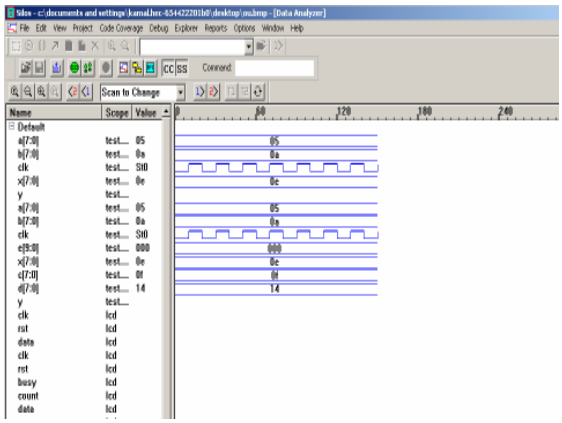


Fig 9 Simulated Waveform for Trapzoidal function

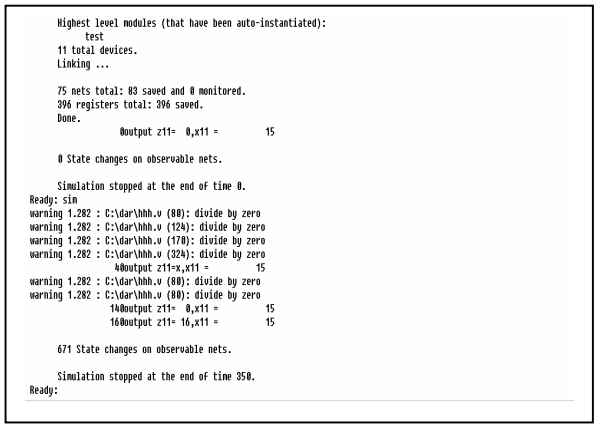


Fig 10 Simulated result for Fuzzy processor

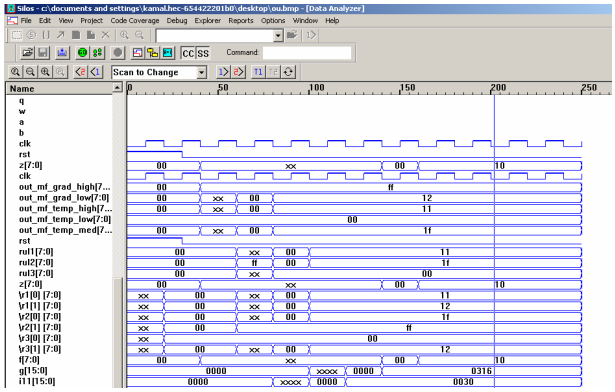


Fig 11 Simulated waveform of Fuzzy processor

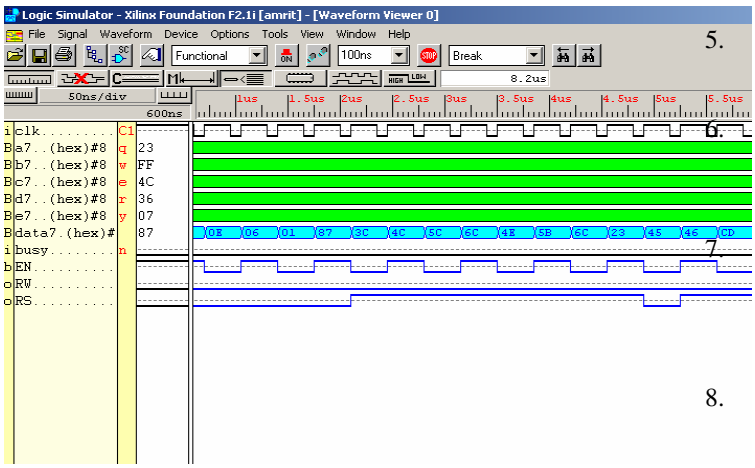


Fig 12 Simulated waveform of LCD Display

CONCLUSIONS

This paper presented implementation of a Designed & developed embedded controller (rapid nickel cadmium battery charger) based on Fuzzy logic with LCD interfacing. The key concept behind a rapid battery charger was the observation that a maximum of charging current upto 8C can be applied to a 2AA Ni-Cd battery upto about the first 5 minutes of charging. Further as the temperature rises, the charging current is controlled based upon the temperature and temperature gradient of the cell. The charging has to be controlled very carefully as the temperature rises because the battery gets damaged at around 50^{o} C.

References:

1. Alfredo Sanz, “Analog Implementation of Fuzzy Controller”, *Proceedings of Third IEEE International Conference on Fuzzy Systems* (FUZZ-IEEE' 94), Orlando, FL, June 26-29, 1994, pp.279-283.
2. Giuseppe Ascia, Vincenzo Catania and Marco Russo, “VLSI Hardware Architecture for complex Fuzzy Systems”, IEEE Trans on Fuzzy systems, Vol 7 N0.5, Oct, 1999. pp 553-570.
3. ALESSANDRA Costa, Allessandro de Gloria and Mauro Olivieri, “Hardware design of Asynchronous fuzzy controllers”, IEEE Transactions on Fuzzy Systems Vol.4, N0.3 Aug 1996, pp 328-338.
4. LX Wang and JM Mendel, “Generating fuzzy rules by learning from examples”, IEEE Transactions on Systems, Man and Cybernetics, Vol.22, N0. 6, Nov-Dec, 1992.pp 1414-1427.
5. JM Mendel, “Fuzzy logic systems for engineering: A Tutorial”, Proceedings of IEEE, Vol.83, No3, Mar, 1995.pp345-377.
6. CC Lee, "Fuzzy Logic in Control Systems: Fuzzy Logic Controllers - Part I", IEEE on Systems, Man and Cybernetics, Vol 20, No. 2, Mar/April, 1990, pp 404-417.
7. CC Lee, "Fuzzy Logic in Control Systems: Fuzzy Logic Controllers - Part II", IEEE on Systems, Man and Cybernetics, Vol 20, No. 2, Mar/April, 1990, pp 419-435. George C. Mouzouris,
8. Marek J. Patyra, Janos L Grantner and Kirby Koster," Digital Fuzzy Logic Controller: Design and Implementation", IEEE Transactions on Fuzzy Systems, Vol 4, No. 4, Nov., 1996. pp 439-459.

9. J. Patyra, D. M. Mlynek (Editors), "Fuzzy Logic : Implementation and Applications", John Wiley & Sons Ltd. and B. G. Teubner, 1996.
10. GM Abdelnour, CH Chang, JY Cheung, “Design of fuzzy logic controller using input and output mapping”, IEEE Transactions on Systems, Man and Cybernetics, Vol.21, N0. 5, Nov, 1991
11. Drinkov, Hellendoorn and Reinfrank, An Introduction to Fuzzy Logic Control.
12. Arun Khosla, “Design and Development of RFC-10:A Fuzzy Logic Based Rapid Battery Charger”, M.Tech. Thesis, Kurukshetra University, 1997.
13. Johan Yen, Reza Langari, “Fuzzy Logic: Intelligence, Control and Information” Prentice Hall International , Inc. 1999.
14. Bart Kosko, “Fuzzy Engineering” Prentice Hall International , Inc. 1997.

SHAPE BASED HYBRID REGION FILLING ALGORITHM

A.Kethsy Prabavathy
II M.E CSE
Department of Computer Science
and Engineering
K.S.Rangasamy College of Technology,
KSR Kalvi Nagar,
Thiruchengode-637 209
Tamil Nadu, India.
91-9443948498
kethsyprabavathy@yahoo.co.in

ABSTRACT

Texture synthesis has a variety of applications in computer vision, graphics, and image processing. An important motivation for texture synthesis comes from texture mapping. Texture images usually come from scanned photographs, and the available photographs may be too small to cover the entire object surface. In this situation, a simple tilting will introduce unacceptable artifacts in the forms of visible repetition and seams. Texture synthesis solves this problem by generating textures of the desired sizes. With the need for various advancement in the functioning of the digital camera, region-filling after object removal from a digital photograph becomes a key issue. This problem is defined as how to "guess" the Lacuna region after removal of an object by replicating a part from the remainder of the whole image with visually plausible quality. The proposed system, develop a hybrid region-filling algorithm composed of a texture synthesis technique and an efficient interpolation method with a refinement approach: 1) The "sub patch texture synthesis technique" can synthesize the Lacuna region with significant accuracy, 2) The "weighted interpolation method" is applied to reduce computation time and 3) The "artifact detection mechanism" integrates the Kirsch edge detector and color ratio gradients to detect the artifact blocks in the filled region after the first pass of filling the Lacuna region when the result may not be satisfactory. The experimental results show that our proposed algorithm can achieve a better performance than previous methods. Particularly, the regular computation of our proposed algorithm is more suitable for implementation of the hardware in a digital camera.

Categories and Subject Descriptors

I.4.1 Digitization and Image Capture [Camera calibration]

General Terms

Algorithms, Documentation.

Keywords

Texture synthesis, Lacuna region, hybrid region-filling, weighted interpolation.

INTRODUCTION

The modification of images in a way that is non-detectable for an observer who does not know the original image is a practice as old as artistic creation itself. Medieval artwork started to be restored as early as the Renaissance, the motives being often as much to bring medieval pictures "up to date" as to fill in any gaps. This practice is called retouching or inpainting. The object of inpainting is to reconstitute the missing or damaged portions of the work, in order to make it more legible and to restore its unity.

Region filling is filling the Lacuna region left after object removal with reasonable feeling for human visual system. The input for our proposed region-filling algorithm is an image with a Lacuna region (the region left after removal of the object,) and our proposed algorithm can infer the Lacuna region from the remaining part of the image (the source region,). The "valid" pixels of the source region serve as examples for filling the Lacuna region. Texture synthesis algorithms generate a large area of similar texture from sample texture or fill the lost region with input texture. Image inpainting algorithms are used to repair the scratches or cracks of photographs and paintings. Generally speaking, texture synthesis is applied to problems of single texture and image inpainting is used in general images with multiple textures. Here, we propose a novel texture synthesis technique of our hybrid region-filling algorithm.

2. LITERATURE REVIEW

Classical image denoising algorithms do not apply to image inpainting. In common image enhancement applications, the pixels contain both information about the real data and the noise (e.g., image plus noise for additive noise), while in image inpainting, there is no significant information in the region to be inpainted. The information is mainly in the regions surrounding the areas to be inpainted.

Mainly three groups of works can be found in the literature related to digital inpainting. The first one deal with the

restoration of films, the second one is related to texture synthesis, and the third one, a significantly less studied class though very influential to the work here presented, is related to disocclusion.autoregressive models are used to interpolate losses in films from adjacent frames. The basic idea is to copy into the gap the right pixels from neighboring frames. The technique can not be applied to still images or to films where the regions to be inpainted span many frames. On the other hand, the algorithm mainly deals with texture synthesis (and not with structured background), and requires the user to select the texture to be copied into the region to be inpainted. For images where the region to be replaced covers several different structures, the user would need to go through the tremendous work of segmenting them and searching corresponding replacements throughout the picture.

The algorithm performs inpainting by joining with geodesic curves the points of the isophotes (lines of equal gray values) arriving at the boundary of the region to be inpainted. As reported by the authors, the regions to be inpainted are limited to having simple topology, e.g., holes are not allowed. In addition, the angle with which the level lines arrive at the boundary of the inpainted region is not (well) preserved: the algorithm uses straight lines to join equal gray value pixels.

2.1 Texture Synthesis

In previous works of texture synthesis, the techniques are classified mainly into three categories. The first category is the synthesis of texture by simulating the physical generation process. The second category is the derivation of a parametric model by analyzing the input texture and synthesizing the output texture. Heeger and Bergen modeled textures by matching marginal histograms of image pyramids. Portilla and Simoncelli presented the application of joint wavelet coefficients to synthesize textures. However, these approaches are not able to capture the local features of texture.

The third category of texture synthesis algorithms is the generation of the output texture by reproducing the sample texture. This is usually accomplished by synthesizing the pixels with the highest level of similarity to the Lacuna region. The pixel-based texture synthesis methods synthesize one pixel at a time. Efros and Leung proposed a nonparametric sampling method based on Markov random fields (MRF). The new pixel is synthesized by choosing one pixel at random from the most similar block of sample texture.

For stochastic texture, pixel-based texture synthesis algorithms always generate better results. However, the serious problem with pixel-based texture synthesis algorithms is huge computation time. These algorithms also fail to reconstruct the global features of structural texture. In order to maintain the global features of texture, patch-based texture synthesis algorithms have been proposed.

Efros and Freeman proposed a simple and effective method by first placing the patches from the input texture randomly as overlapping tiles. The mismatching across the boundary of patches is fixed by the minimal error cut, which is analogous to dynamic programming. Liang et al. presented a real-time patch-based texture synthesis algorithm using samples blending. These methods can produce texture well by patch-based sampling. However, the patch boundary problem is the main issue of patch-based texture synthesis algorithms. This approach, however, is still not suitable for the general image.

2.2 Image Inpainting

Considering the image inpainting algorithms the partial differential equation (PDE) used in image inpainting was introduced first by Bertalmio et al. The source region is diffused inwardly from the boundary of the Lacuna region. The propagation direction is estimated as the minimal spatial change by orthogonal to isophote direction of the Laplacian.

Bertalmio et al. proposed a state-of-the-art algorithm for integrating texture synthesis and image inpainting. The texture synthesis part of the algorithm synthesizes one pixel at a time from the right pixel of the matched block. However, the image inpainting part of the algorithm always produces a blurring effect in the filled region and nontrivial parameters needed to be adjusted manually.

Recently, Criminisi et al. proposed an exemplar-based region-filling algorithm after removing large objects from a digital image. The algorithm iteratively fills the lost region by block-based sampling. The priority of synthesizing order is decided by the pixel number in the source region and the dot product between the isophote direction and the normal vector. The blocks for each pixel on the boundary of the Lacuna region are compared with the sample blocks of the source region to obtain the filled blocks with the highest level of similarity. A priority updating step is needed for each step of the filling process.

Previous algorithms provide only a single method for all kinds of input images. An extra color space transform is also needed in previous algorithms. The thesis proposed a complete algorithm i.e., two independent methods are applied to fill the Lacuna region by analyzing the texture characteristic and the artifact detection mechanism is employed to detect the artifact blocks.

3. METHODOLOGY

3.1 Color Texture Distribution Analysis

Color texture distribution analysis is performed to choose the proper method for region-filling the Lacuna region. The real-world images always include several textures and horizontal oriented background. To avoid unnecessary computation time, the proposed methodology uses a weighted interpolation method in these regions. Using color texture distribution analysis the proposed methodology, decide whether the sub-patch texture synthesis technique or the weighted interpolation method should be applied.

The color texture distribution analysis can distinguish the color distribution and spatial structure near the Lacuna region. By observation of the texture characteristic of the surrounding valid pixels, the proposed methodology chooses the processing method to use for filling the Lacuna region. The valid pixels around the Lacuna region are separated into the inhomogeneous texture region and the homogeneous texture region by computing the characteristic value. In order to maintain the quality of output image, some homogeneous texture may be classified into inhomogeneous texture in our

definition. The color texture distribution analysis comprises of computing the characteristic value, marking the classification map and determining the synthesized part and the filled part of the Lacuna region.

3.2 Subpatch Texture Synthesis Technique

The subpatch texture synthesis technique is applied for the synthesized part of the Lacuna region. Compared with the traditional pixel-based synthesis algorithm, the main feature of our subpatch texture synthesis technique is pasting a subpatch (line) each time with great regularity and efficiency. The subpatch texture synthesis technique comprises of three parts i.e., defining the notations, determining L and U-shaped orientation by steerable pyramid decomposition, and adaptively searching the best candidate patch.

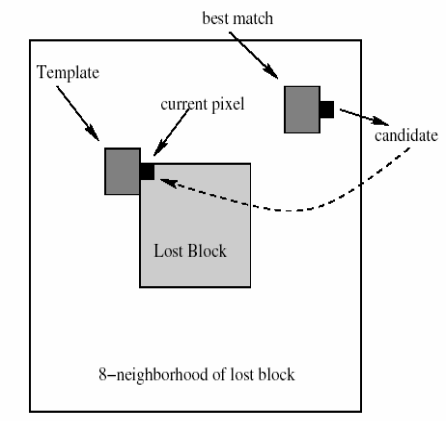


Figure 1: Basic texture synthesis procedure

3.3 Weighted Interpolation Method

For the filled part of the Lacuna region, the weighted interpolation method is used to infer the Lacuna region. It assumes that the background of the real-world image consists mostly of horizontal textures. The weighted interpolation method uses the pixels in the same row of source region to infer the pixels of the Lacuna region.

The color values of the leftmost pixel in the Lacuna region are calculated. If the pixels are not valid pixels of the source region, use the pixels of the source region to pre-estimate the pixel values. The pixels in the two sides are interpolated first to be used as the sample pixels for inner pixels and each is a weighting value. On the other hand, the value of the rightmost pixel in the Lacuna region is calculated in an analogous manner. The weighting values are the Gaussian kernel to maintain the local feature.

3.4 Artifact Detection Mechanism and Resynthesizing

In order to provide a robust quality of output image, if the result is not satisfactory after the first pass of region filling, the user can apply the artifact detection mechanism to the filled region. In previous image inpainting algorithms, there is no solution to detect if the filled region is unlike the original background.

The proposed artifact detection mechanism integrates the Kirsch edge operator and color ratio gradients to detect the artifact blocks in two-stages. At the first stage, the Kirsch candidate blocks are detected by the Kirsch edge operator in a larger block size. Then, the artifact blocks are selected from the Kirsch candidate blocks by color ratio gradients in smaller block size to acquire a more accurate result.

4. SYSTEM MODEL AND EXPERITMENTAL EVALUATION

At first, color texture distribution analysis is performed on the Lacuna region to determine which method is to be used in the Lacuna region. After the first pass of filling the Lacuna region, the result with the filled Lacuna region is generated.

Optionally, if the result is not satisfactory, the artifact detection mechanism will detect the artifact blocks in the filled region and color the artifact blocks in white. After the second pass of region filling, these artifact blocks will be resynthesized by a subpatch texture synthesis technique to generate the final output image.

The simulation environment is produced on a 2.6 GHz Pentium IV with 1GB of RAM and is implemented in MatLab 7.0. The proposed algorithm consumes 12 minutes to fill the Lacuna region successfully. Another example is used to evaluate the performance of edge connection in a natural scene. The rainbow is reconstructed to connect to the other side. Comparing with tensor voting, the red chair is removed and the image is filled in. The original image and the image after object removal. The grove is generated more naturally due to the growing direction.

The classification map from color texture distribution analysis is taken for analysis. The left region after removal of the lighthouse is filled by our proposed algorithm. The artifact blocks are detected by color ratio gradients.

After implementation of the artifact detection mechanism, the artifact blocks are resynthesized by the subpatch texture synthesis technique. The experiments apply the artifact detection mechanism to the result from image completion.

The parameters used in the thesis are the neighborhood size of the Gaussian kernel for computing a fast approximation is set at 7, the neighborhood size of the Gaussian kernel for constructing the Laplacian pyramid is set at 9 and the standard deviation is set at 1, to get the output result. The Kirsch candidate blocks are colored in white. Then, the artifact blocks are detected by color ratio gradients. The structure propagation is important to finish the output image after region filling.

To evaluate the robustness of our proposed subpatch texture synthesis technique, the most important issue is edge connection. For the perceptivity of the human sense, edge broken is considered as a failure. In previous papers, the synthesizing order is suggested to dominate the quality of the output image. The edges of the source region must be propagated to connect the image structure.

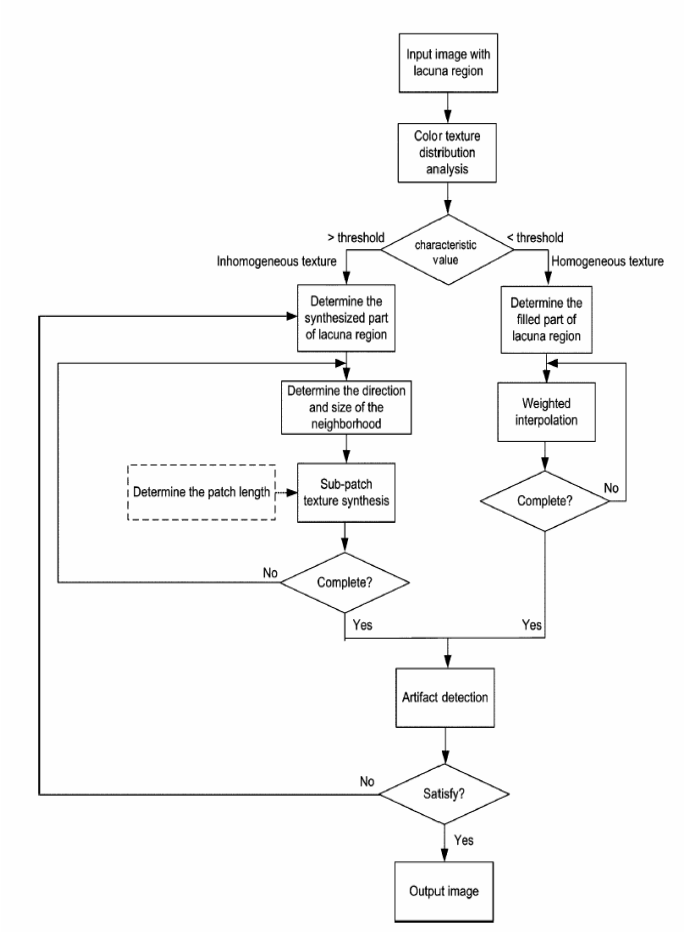


Fig 2: Flowchart of our proposed region-filling algorithm.

For pure texture synthesis, the subpatch texture synthesis technique is useful for pure texture. Our algorithm can generate an output image without an extra color space transform such as RGB to or YUV. The subpatch texture synthesis technique also can be reduced to a pixel-based algorithm by setting the patch length to 1, which is especially useful for structure image inpainting. Besides, the efficiency is defined as the ratio of synthesized pixel number to the sampling pixel number.

5. CONCLUSION

The region-filling algorithm that offers robust results for region filling after object removal from a digital photograph is developed in this thesis. Color texture distribution analysis determines whether the weighted interpolation method or the subpatch texture synthesis technique will be applied. The weighted interpolation method is used to reduce the computation time in searching for the best samples. The proposal can generate plausible visual results without the blurring effect.

The image inpainting that attempts to replicate the basic techniques is used by professional restorators. The basic idea is to smoothly propagate information from the surrounding areas in the isophotes direction. The user needs only to provide the region to be inpainted; the rest is automatically performed by the algorithm in a few minutes. The inpainted images are sharp and without color artifacts. The results can either be adopted as a final restoration or be used to provide an initial point for manual restoration, thereby reducing the total restoration time by orders of magnitude. The ideal algorithm should be able to automatically switch between textured and geometric areas (Shape), and select the best suited technique for each region.

The proposals provide an artifact detection mechanism to detect the artifact blocks and resynthesize the artifact blocks by a subpatch texture synthesis technique. The greater regularity of our algorithm is also suitable to implement the hardware architecture to be embedded in digital camera.

6. REFERENCES

[1] D. J. Heeger and J. R. Bergen, "Pyramid-based texture analysis / synthesis," in Proc. ACM Conf. Computer Graphics, Los Angeles, CA, 1995, vol. 29, pp. 229–233.

[2] J. Portilla and E. P. Simoncelli, "A parametric texture model based on joint statistics of complex wavelet coefficients," Int. J. Comput. Vis., vol. 40, no. 1, pp. 49–71, Oct. 2000.

[3] A. Efros and T. Leung, "Texture synthesis by non-parametric sampling," in Proc. Int. Conf. Computer Vision, Kerkyra, eece, Sep. 1999, pp. 1033–1038.

[4] M. Ashikhmin, "Synthesizing natural textures," in Proc. ACM Symp. Interactive 3D Graphics, Mar. 2001, pp. 217–226.

[5] A. Efros and W. T. Freeman, "Image quilting for texture synthesis and transfer," in Proc. ACM Conf. Computer Graphics, Aug. 2001, pp. 341–346.

[6] L. Liang, C. Liu, Y. Q. Xu, B. Guo, and H. Y. Shum, "Real-time texture synthesis by patch-based sampling," ACM Trans. Graph., vol. 20, pp. 127–150, 2001.

[7] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in Proc. ACM Conf. Computer Graphics, Jul. 2000, pp. 417–424.

[8] M. Bertalmio, L. Vese, G. Sapiro, and S. Osher, "Simultaneous structure and texture image inpainting," IEEE Trans. Image Process., vol. 12, no. 8, pp. 882–889, Aug. 2003.

A genetic algorithm for producing analogous proteins

Dominik Heider
Department of Experimental Tumorbiology, University of Muenster
Badestrasse 9
D-48149 Muenster, Germany
dominik.heider@uni-muenster.de

ABSTRACT
This paper provides a genetic algorithm for creating proteins, which have the same secondary structure and hydrophobizity characteristics as a given protein. These analogous proteins are examples for non related and non descending proteins, which have similar structures and might have similar functions. It uses binary fighting selection, one-point crossover and point mutations. For our experiments we use the Ypt7 protein sequence of *Saccharomyces cerevisiae*. *In silico* studies using a prototype exhibits a structure similarity over 60%, which can be improved using more significant characteristics of the protein. This work may be the first step to developing well defined recombinant antibodies for cancer therapies.

Keywords
Genetic algorithm, protein structure, protein design

1. INTRODUCTION
Genetic algorithms can be used for optimization problems. The idea is copied from nature, where genetic information in organisms is encoded in their genome, which consists of deoxyribonucleic acid (DNA). DNA is a long chain of four different nucleotides, adenine, cytosine, guanine and thymine. The order of these four nucleotides encode the information, which contains the genetic instructions used in the development and functioning of all organisms. The genetic code, which specifies the amino acid sequence within a protein, is used to read the instructions. A gene is a distinct region which encodes parts or even entire proteins. A genotype of an organism is described by its alleles. An allele is a specific DNA sequence at a gene. Different individuals may have identical alleles, but can also have different alleles. Alleles can be used to exhibit relations between organisms or even individuals.

Genetic algorithms (GAs) are an artificial copy of this constellation. An individual in GAs is the equivalent genotype of an allele. A population is a set of all individuals.

In general, genetic algorithms begin with a random start population. In the next step a new population is created using selection, recombination and mutation. This procedure is repeated several times and ends if a fixed number of iteration is reached or an individual with a specified fitness is found. The fitness is the quality of the individual (solution). For our experiments we use the Ypt7 protein sequence of *Saccharomyces cerevisiae*. Ypt7 is one of the 12 members of the Ypt GTPase family [1]. The mammalian homolog of Ypt7 is Rab7. GTPases are molecular switches in cellular transport processes, which can be either in an active GTP (Guanosine-5'-triphosphate) bounded or in an inactive GDP (Guanosine diphosphate) bounded constellation [2].

The Ypt7 GTPase from *S. cerevisiae* is involved in late endosome-to-vacuole transport and vacuole fusion events [3,4].

In nature, analogous characteristics e.g. the wings of birds and insects, frequently appear. These characteristics are non-related with each other, but perform the same task. At protein level, similar examples exist.

Independent evolution of similar proteins as a result of having to adapt to similar environments is called convergent evolution. An example at protein level are the antifreeze proteins [5-6].

The function of a protein is determinded by its structure, so biologists have to search for structure and hydrophibicity similarities instead of primary sequence identity to find functional similitudes.

2. METHODS
2.1 Genotype
Every individual in the population represents an amino acid sequence of a fixed length (Table 1).

amino acid	code	amino acid	code
Alanine	A	Leucine	L
Arginine	R	Lysine	K
Asparagine	N	Methionine	M
Aspartic acid	D	Phenylalanine	F
Cysteine	C	Proline	P
Glutamic acid	E	Serine	S
Glutamine	Q	Threonine	T
Glycine	G	Tryptophan	W
Histidine	H	Tyrosine	Y
Isoleucine	I	Valine	V

Table 1: One letter code for amino acids.

2.2 Start population
The start population contains a fixed number of individuals, whose genotypes are created randomly from the amino acid set.

2.3 Fitness function
The fitness is represented by $\phi \in [0,1]$. The fitness describes the quality of an individual and is expressed via the similarity of the secondary structure (α), the similarity of the hydrophobicity characteristics (β) and the similarity of the amino acid sequence

(δ) [7-9]. The secondary structure prediction was done using PREDATOR with standard settings [8-9]. The hydrophobicity plots were done using the algorithm of Kyte and Doolittle [7].

$$\phi = \text{mean}(\alpha,\beta) / \delta$$

The similarity can be calculated by exact pattern matching or by using sequence alignments [10,11].

2.4 Selection

50% of the individuals of a generation are transferred directly to the next generation. They are selected by binary fighting selection which means that two individuals are selected randomly and their fitness is compared and the one with the better fitness is transferred to the next generation without any modification.

2.5 Genetic operations

The genetic algorithm uses two forms of genetic operations. The first one is recombination and the second one is mutation.

2.5.1 Recombination

50% of individuals in a generation are created by recombination, which is performed by one-point-crossover. A random crossover point (i) is randomly choosen and two individuals (A, B) are randomly selected. If the selected individuals have a fitness greater than a random number, two new individuals (C, D) are created with

C = A[0..i]B[i..n]

D = B[0..i]A[i..n], where n is the length of an individual

If the fitness is lower than a random number, two new individuals are selected. This procedure prefers individuals with a better fitness.

2.5.2 Mutation

After creating the next generation, point mutations occur randomly. A point mutation changes the existing amino acid at a random position in another random amino acid. Mutations lead to a greater search space.

3. RESULTS

The number of individuals per generation was set to 36. The prototype of the genetic algorithm has a fitness function which differs slightly from that described before. The fitness is calculated solely based upon the similarity of the secondary structure prediction. 500 mutations are made with every new generation.

Tests were performed using 10 and 20 generations.

Table 2 shows the mean value (mean), standard deviation (δ) and maximum value (max) of the best individual fitness using 10 and 20 generations. The best resulting amino acid sequence generated by the genetic algorithm using 10 or 20 generations have a fitness of 0.577 and 0.601, respectively. Multiple sequence alignments using ClustalW show no significant identical parts of the amino acid sequences (Figure 1) [12,13].

generations	mean	δ	max
10	0.52	0.04	0.577
20	0.52	0.05	0.601

Table 2: The mean value (mean), standard deviation (δ) and maximum value (max) of the best individual fitness using 10 and 20 generations.

The secondary structure prediction and the hydrophobicity plots of the Ypt7 sequence and the sequences produced by the genetic algorithm exhibit, that there is a high similarity between the characteristics of Ypt7 and the amino acid sequence produced within 20 generations (Figure 2, 3) [7-9].

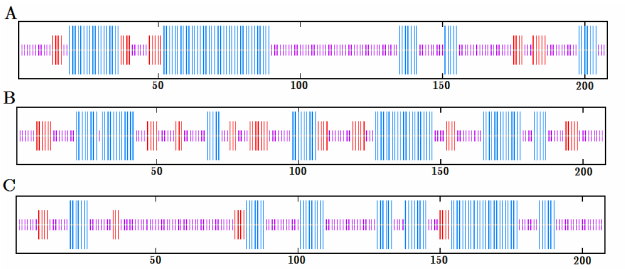


Figure 2: A: Structure after 10 generations of optimization; B: structure of Ypt7; C: structure after 20 generations of optimization. Alpha helices are blue, beta sheets are red and random coils are purple.

4. CONCLUSION

A genetic algorithm with a well defined fitness function can be used to produce proteins with a high similar secondary structure and hydrophobicity characteristics to a given protein sequence.

The algorithm described in this paper is the first step producing artificial proteins which can be used in future for creating well designed recombinant antibodies for cancer therapies using more detailed fitness functions including tertiary structure predictions and protein interaction predictions. The confirmation of these *in silico* results will be task of further examinations.

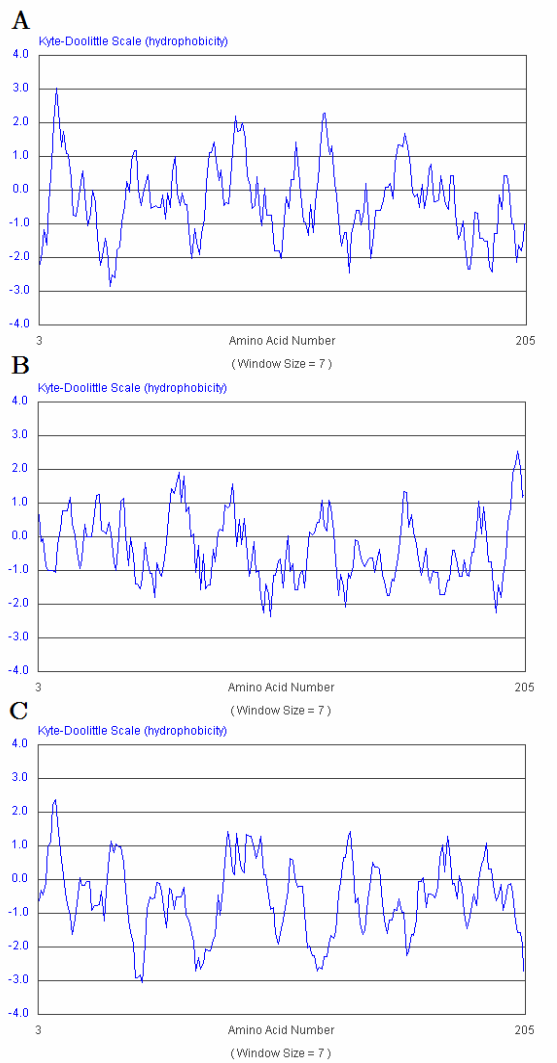


Figure 3: Hydrophobicity plots. A: Ypt7, B:sequence produced within 10 generations, C: sequence produced within 20 generations. A score greater than 0 indicates a hydrophobic region. A window size of 7 was used [7].

5. REFERENCES

[1] Neuhaus J, Schulze U, Barnekow A, Kail M: Identi_cation of a novel small GTPase in yeast. Europ. J. Cell Biol. 2007, 86:58.

[2] Watzke A, Brunsveld L, Durek T, Alexandrov K, Rak A, Goody R, Waldmann H: Chemical biology of protein lipidation: semi-synthesis and structure elucidation of prenylated RabGTPases. Org Biomol Chem. 2005, 3:1157–1164.

[3] Wichmann H, et al: Endocytosis in yeast: evidence for the involvement of a small GTP-binding protein (Ypt7p). Cell 1992, 71:1131–1142.

[4] Schimoller F, Riezmann H: Involvement of Ypt7p, a small GTPase, in traffic from late endosome to the vacuole in yeast. J Cell Sci 1993, 106:823–830.

[5] Pertaya N. et al (2007): Basal Plane Affinity of an Insect Antifreeze Protein. [http://meetings.aps.org/link/BAPS.2007.MAR.J35.8]

[6] Pertaya N. et al (2007): Fluorescence microscopy evidence for quasi-permanent attachment of antifreeze proteins to ice surfaces. Biophys J 2007

[7] Kyte J, Doolittle RF: A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 1982, 157:105–132.

[8] Frishman D, Argos P: Incorporation of long-distance interactions into a secondary structure prediction algorith. Protein Engineering 1996, 9:133–142.

[9] Frishman D, Argos P: 75 percent accuracy in protein secondary structure prediction. Proteins 1997, 27:329–335.

[10] Needleman BS, Wunsch CD: A general method applicable to the search for similarities in the amino acid sequence of two proteins. J. Mol. Biol. 1970, 48:443–453.Tavel, P. 2007 Modeling and Simulation Design. AK Peters Ltd.

[11] Smith TF, Waterman MS: Identification of common molecular subsequences. J. Mol. Biol. 1981, 147:195–197.

[12] Thompson JD, Higgins DG, Gibson TJ: CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994, 22:4673–4680.

[13] European Bioinformatics Institute ClustalW [http://www.ebi.ac.uk].

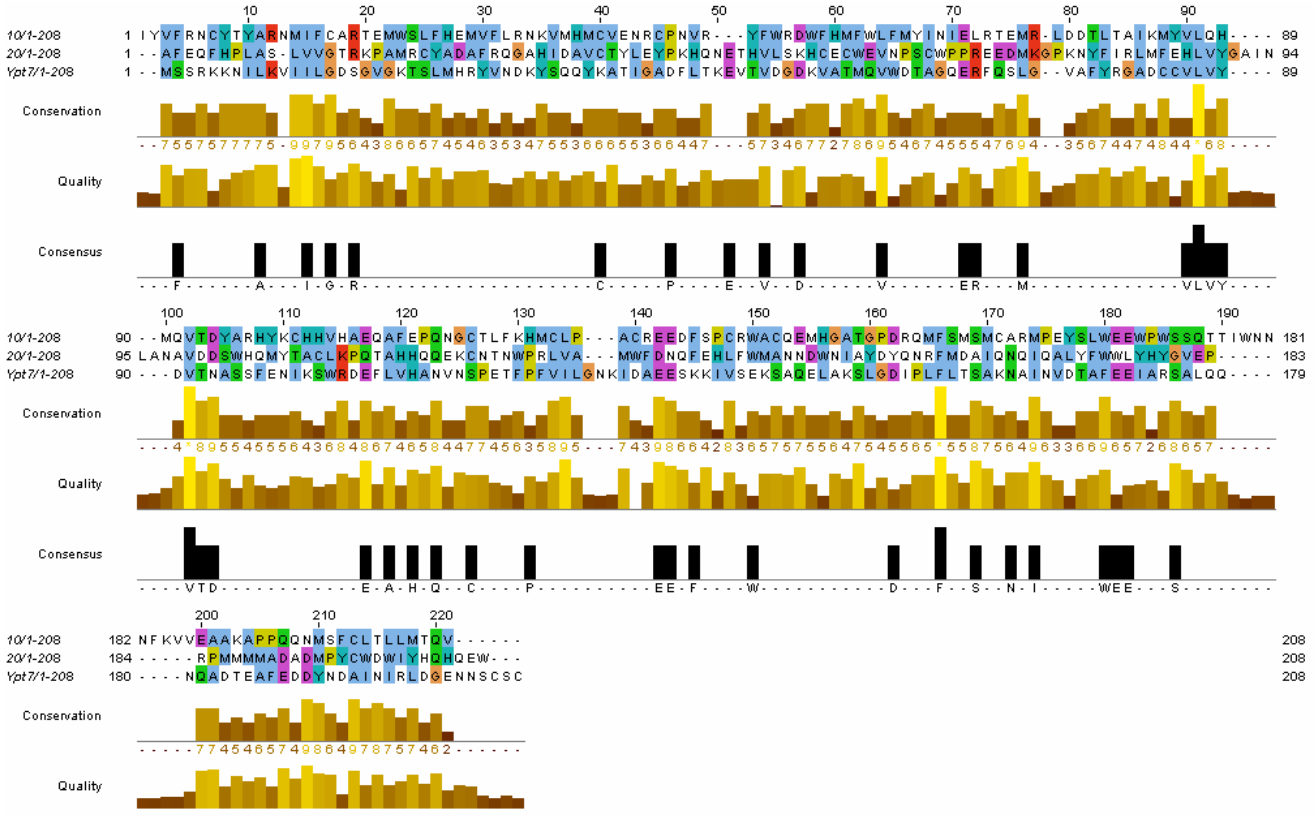


Figure 1: Multiple sequence alignment of the Ypt7 sequence and the sequence of the two analogous amino acid sequences produced by 10 or 20 generations [12, 13]. The consensus exhibits homolog parts of the sequences.