IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 2, March 2011
ISSN (Online): 1694-0814
www.IJCSI.org


# BW Trained HMM based Aerial Image Segmentation

R.Rajasree[1], J. Nalini[2], S.C.Ramesh[3]

[1] Communication Systems,PSN College Of Engineering and Technology, Melathedioor, Tirunelveli Tamilnadu , India

[2] Electronics and Communication,PSN College Of Engineering and Technology, Melathedioor, Tirunelveli Tamilnadu , India

[3] Avionics, PSN College Of Engineering and Technology, Melathedioor, Tirunelveli Tamilnadu , India

## Abstract

Image segmentation is an essential preprocessing tread in a complicated and composite image dealing out algorithm. In segmenting arial image the expenditure of misclassification could depend on the factual group of pupils. In this paper, aggravated by modern advances in contraption erudition conjecture, I introduce a modus operandi to make light of the misclassification expenditure with class-dependent expenditure. The procedure assumes the hidden Markov model (HMM) which has been popularly used for image segmentation in recent years. We represent all feasible HMM based segmenters (or classifiers) as a set of points in the beneficiary operating characteristic (ROC) space. optimizing HMM parameters is still an important and challenging work in automatic image segmentation research area. Usually the Baum-Welch (B-W) Algorithm is used to calculate the HMM model parameters. However, the B-W algorithm uses an initial random guess of the parameters, therefore after convergence the output tends to be close to this initial value of the algorithm, which is not necessarily the global optimum of the model parameters. In this project, a Adaptive Baum-Welch (GA-BW) is proposed.

*Key terms*—Convex hull, hidden Markov models, image segmentation, ROC convex analysis, ROC curve, Genetic Algorithm; HMM training; Baum-Welch algorithm..

## 1.INTRODUCTION

IMAGE segmentation extracts explicit information about continuous, and it allows human observers to understand images clearly by focusing specific regions of interest. For this reason, it is often used as an initial procedure to simplify a sophisticated and complex image processing system. Segmentation often requires smooth boundary between regions for different classes, and hidden Markov model (HMM) is possibly one of the most popular models for it. The HMM assumes that the true hidden class has Markovian dependency and, thus, has smooth boundary between segmented regions.

The popularly used Markov models have two parameters, which we denote by $\alpha$ and $\beta$ where indicate the popularity of each class over images and the other $\boldsymbol{\beta}$ indicates the strength of spatial coherence. The parameters are estimated in segmentation procedures or pre-decided by an expert. In real example, the cost of misclassification can depend on the classes.

For example, in cancer diagnosis, misclassifying cells normal cells pays much higher cost than normal cells to tumor cells; or, in segmenting and detecting military targets, mistakenly detecting targets to non targets may cost more than the other type of misclassification. However, all existing segmentation procedures do not take into account the cost of misclassification, particularly the unequal cost that depends on the classes.

But In this existing System to segment aerial images with two classes, “Artificial region” and “natural region,” which could be used for target recognition and tracking. Here, the cost of misclassifying targets (or Artificial region) is higher than that of misclassifying non-targets (or natural region) to targets. However, the B-W algorithm uses an initial random guess of the parameters, therefore after convergence the output tends to be close to this initial value of the algorithm, which is not necessarily the global optimum of the model parameters. In this project, a adaptive Baum-Welch (BW) is proposed.

## 2. REVIEW OF ROC

The ROC convex analysis draws great attention in the ma-chine learning society. In a two-class problem (positive and negative class), the ROC curve (or set) is the plot of the probability of false positive decision (false positive rate, FPR) and that of true positive decision (true positive rate, TPR). The ROC convex hull analysis finds an optimum point in an ROC space to minimize the misclassification cost of classifier, which is defined on which costs s of classifiers are the same.

$$\text{Cost}(\iota) = \text{FPR}(\iota) + \text{FNR}(\iota) \quad (1)$$

Then, the ROC convex analysis finds the optimal classifier as the classifier whose (FPR, TPR) pair is the tangential point between these iso-cost lines and the convex hull of the ROC curve.

The ROC convex analysis requires the entire set of feasible classifiers. In the HMM-based segmentation, the model has two unknown parameters which are estimated in segmenting the image .Fixing these two parameters, not estimating, provides the set of classifiers which are indexed by the ROC set is the set Where FNR(ι) implies false negative rate which is equal to 1-TNR(ι). Once misclassification cost is given, we have a family

$$\alpha = (\text{Unit cost of FNR / Unit cost of FPR}) \quad (2)$$

on which costs of classifiers are the same. Then, the ROC convex analysis finds the optimal classifier as the classifier who's (FPR, TPR) pair is the tangential point

## 3. HMM TRANING SYSTEMS

The model parameters of an HMM can be expressed as a set of three elements: λ = {A, B, n } [11]. Where:

- A= { aij } is the state transition probability matrix, each element aij represent the probability that the HMM model will transit from state i to state j. Elements of matrix A must satisfy the next two conditions:

$$aij \geq: 0 \quad \text{where } 1\leq i, j \leq 3 \quad (3)$$

$$\Sigma^3_{j=0}\, bij = 1 \quad \text{where} 1\leq i,j\leq 3 \quad (4)$$

- B ={bij ( k )} is the observation probability matrix, such that bij is the probability that the observation Ok has been generated by state i. Elements of matrix B must satisfy the next two conditions:

$$bij \geq: 0 \quad \text{where } 1\leq i, j \leq 3 \quad (5)$$

$$\Sigma^3_{j=0}\, bij=1 \quad \text{where} 1\leq i,j\leq 3 \quad (6)$$

- Π = {πi} is the initial state distribution vector, and every πi express the probability that the HMM chain will start at state i. Elements of vector n must satisfy the next two conditions:

$$\Pi i \geq 0 \quad \text{where } 1\leq I\leq 3 \quad (7)$$

$$\Sigma^3_{j=0}\, \Pi i = 1 \quad (8)$$

HMM Training is the process of HMM parameter calculation. This is shown in figure 1, the training tools use the speech data and their transcriptions to estimate the HMM parameters then the recognizer will use these HMMs to classify the unknown speech utterances

## 4. HMM TRAINING USING B-W ALGORITHM

The B-W algorithm provided by MixtGaussian toolbox has been used to train the HMM. After an initial guess of the HMM parameters is made, the B-W algorithm is run for 20 iterations to get more accurate parameters. As result we get a continuous density mixture Gaussian HMM. Finally transcriptions of unknown speech utterances will be made by the recognizer module to determine how accurate are the HMM's parameters.

## 5. Modified HMM using Baum Welch

.The algorithm has two steps:

1) Calculating the forward probability and the backward probability for each HMM state.

2) On the basis of this, determining the frequency of the transition-emission pair values and dividing it by the probability of the entire string . This amounts to calculating the expected count of the particular transition-emission pair. Each time a particular transition is found, the value of the quotient of the transition divided by the probability of the entire string goes up, and this value can then be made the new value of the transition.

### 5.1 Baum-Welch Algorithm

This method can be derived using simple ``occurrence counting'' arguments or using calculus to maximize the auxiliary quantity

$$Q(\lambda, \lambda') = \sum_q p\{Q/O\} \,\mathrm{Log}[P\{O, q, \lambda'\} \quad (9)$$

Over λ'[]. A special feature of the algorithm is the guaranteed convergence. To describe the Baum Welch Algorithm also known as forward backward algorithm .We need to define two more auxiliary variables, in addition to the forward and backward variables defined in a previous section. These variables can however be expressed in terms of the forward and backward variables.

$$\Sigma^3_{j=0}\, \Pi i = 1 \quad (10)$$

First one of those variables is defined as the probability of being in state *i* at *t*=*t* and in state *j* at *t*=*t*+1. Formally

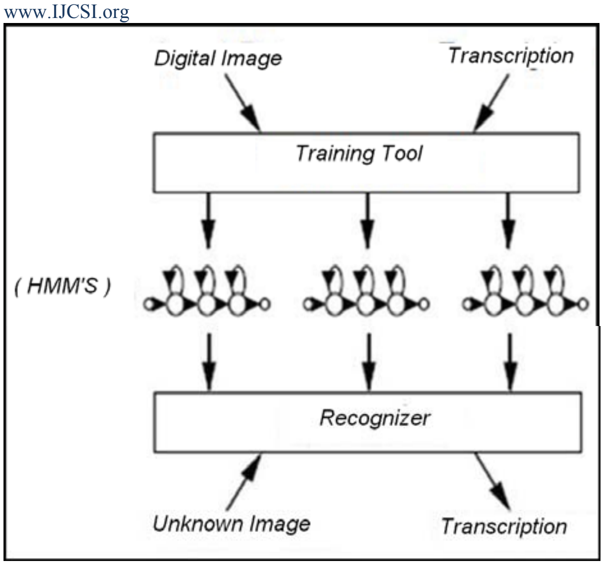


Figure: 1 HMM training process[1]

,

$$\epsilon_t = P\ \{q_t = i,\ q_{t+1} = j/O\ ,\ \lambda\} \quad (11)$$

This is the same as,

$$\epsilon_t = [P\ \{q_t = i,\ q_{t+1} = j/O\ ,\lambda\}]\ /\ P\ \{O/\lambda\} \quad (12)$$

Using forward and backward variables this can be expressed as,

$$\epsilon_t = [\ \alpha_t(i)a_{ij}\beta_t + 1(j)b_j(a_{t+1})\ ]\ /\ \sum_{i=n}^{N}\sum_{j=n}^{N}[\ \alpha_t(i)a_{ij}\beta_t + 1(j)b_j(a_{t+1})\ ] \quad (13)$$

The second variable is the a posteriori probability,

$$\Gamma_i(i) = P\ \{q_t = i|O,\ \lambda\} \quad (14)$$

that is the probability of being in state *i* at *t*=*t*, given the observation sequence and the model. In forward and backward variables this can be expressed by,

$$\gamma_i\ = [\alpha_i\beta_t(i)]\ /\ \sum_{i=1}^{N}[\alpha_i\beta_t(i)] \quad (15)$$

One can see that the relationship between $\gamma_i$ and $\epsilon_t$ (i,j) is given. Now it is possible to describe the Baum-Welch learning process, where parameters of the HMM is updated in such a way to maximize the quantity, P {o|λ}. Assuming a starting model λ=(A,B,π),we calculate the α's and β's using the recursions 5 and2, and then Є's and β's using 12 and15. Next step is to update the HMM parameters according to equations 16 to18, known as reestimation formulas.

$$\Pi'_i = \gamma_i\ (i)\ , \qquad 1 \le\ i \le N \quad (16)$$

$$a'_{ij} = \sum_{t=1}^{T-1}\epsilon_t\ (i\ ,\ j)\ /\ \sum_{t=1}^{T-1}\gamma_t(i) \quad 1 \le\ i \le N \quad (17)$$

These reestimation formulas can easily be modified to deal with the continuous density case too. Convergence It can be proven that if current estimate is replaced by these new estimates then the likelihood of the data will not decrease (i.e. will increase unless already at a local maxima/critical point). See Durbin, Section 11.6 for discussion of avoiding local maxima and other typical pitfalls with this algorithm.

## 6. HMM TRAINING USING HYBRID GA-BW ALGORITHM

In MixtGaussian toolbox implementation a 3 hidden state continuous density mixture Gaussian HMM with 105 observation symbols have been used, this configuration can well describe the speech utterance. Thus, the HMM model parameters is Ie = {A, B, n }, where A is a 3 by 3 transition probability matrix, B a 3 by 105 matrix, n is the initial probability of states vector of size 3.

6.1 Encoding Method

It is vital to find a genetic representation of the model parameters before applying GA to solve the optimization problem. In genetics, chromosomes are comprised by a set of basic elements called genes, in our case the elementary information is the elements of every probability matrixes A, B and n.

We choose to concatenate the rows of each matrix in the model parameters I.e., thus the chromosomes will be represented as an array of real numbers as shown in Fig. 2.

6.2. Hybrid GA-BW

We present a hybrid GA-BW algorithm due to slow convergence and high computational power needed by classical GA, especially when the generated chromosomes cannot satisfy the conditions (13), (14), . . . ,(18). For this reason, not included in the offspring and they are replaced by new chromosomes.

6.3. Fitness Function

The fitness function is an evaluation mechanism of them chromosome; a higher fitness value reflects the chances of the chromosome to be chosen in the next generation. The log likelihood [1] has been used, and it represents the

ISSN (Online): 1694-0814
www.IJCSI.org

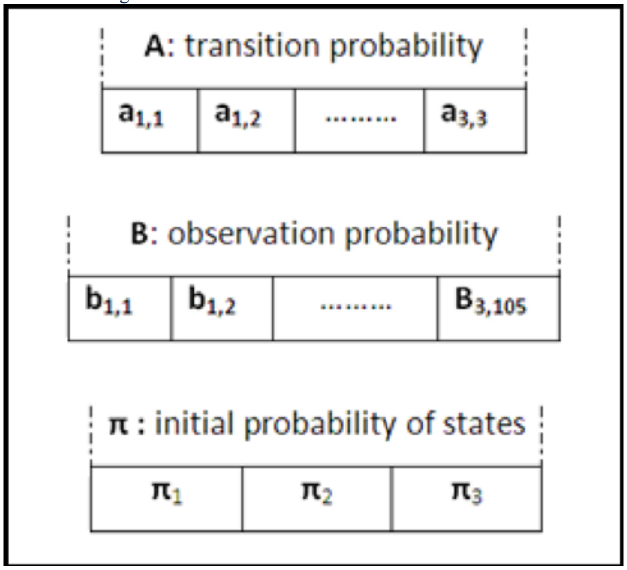


Figure 2. Chromosome encoding

probability that the training observation utterances have been generated by the current model parameters and it is a function of the following form [2]:

1) Apply B-W algorithm to generate the initial Population P(O), where P(O) = {C], C2, ... , CN}, and Cj is one chromosome.

2) Calculate the fitness function F(Cj) of every chromosome Cj within the actual population P(t).

3) Select a few chromosomes for the intermediate population P'(t).

4) Apply crossover to some chromosomes P'(t).

5) Apply mutation to few chromosomes in P'(t)

6) Apply three iterations of B-W algorithm to the population P(t) for each ten generations.

7) t = t+l; if not convergence, then go to: 2).

Fig 3: Steps followed in crossover and mutation using BW

### 6.4. Mutation

Mutation selects randomly few chromosomes and alters some genes to produce new chromosomes. The "Mutationadaptfeasible" function of GA toolbox has been used to satisfy the conditions (3), (4) . . . (8).

### 6.5 Selection and Crossover

Selection mimics the survival of fittest mechanism seen in Nature. Chromosomes with higher fitness values have a greater probability to survive in the succeeding generations. Then some elements from the population pool will be selected to apply crossover. Portions of genes will be exchanged between each Couple of chromosomes. The couple of chromosomes. The default GA toolbox selection and crossover functions have been used in the implementation.

## 7. ROC CONVEX HULL ANALYSIS

In a binary classifier, the ROC curve (or space) plots two accuracy measures of a classifier, FPR and TPR. Suppose we use a continuous classification score X to diagnose a certain disease, and we classify a subject into a disease (positive) group if his/her score is higher than a given critical point; otherwise, we classify him/her into non disease group (negative) group.

The optimal ROC curve is the one produced by the classifiers which has the maximum TPR given FPR. The optimal ROC curve has several geometrical properties including Convexity .Suppose it is not convex on an interval [a, b] belongs to [0, 1] ,where a and b correspond to critical values $c_1$ and $c_2$ in the way that the FPR at $c_1$ and $c_2$ is a and b respectively. Then, for the diagnosis with critical values C Є [$c_1$, $c_2$] we find a better diagnostic system (classifier) which has the same FPR but higher TPR by randomly choosing between two diagnoses with critical values $c_1$ and $c_2$ . Thus, the convex hull of the observed ROC curves represents the ROC curve of the set of potentially optimal classifiers. We let D in ROC space be the set of (FPR, TPR)s of all classifiers we consider, and D'be the set of their random mixtures. Then, D' is the smallest convex region which contains . The goal of this letter is to find the classifier which minimizes the cost (1) among classifiers in . Since FNR (ι) = 1- TPR (ι), Cost (ι) = α FPR(ι) + ( TPR(ι)) and

$$TPR(\iota) = 1 - Cost(\iota) + \alpha\, FPR(\iota) \quad (18)$$

We represent the cost functions Cost(ι)s in the ROC space. Then, the intercept of the iso-cost line (2) is 1- Cost(ι)and, as in Fig.3, the point to minimize the cost is the tangential one between the line with slope α and the convex area D'.

In summary, proposed procedure to find the cost-effective classifier has the following steps:

S1) We define a set of classifiers, say c , to be considered. In case of HMM-based segmentation, this becomes a set of classifiers corresponding to each of two parameters using

S2) We compute FPRs and TPRs of classifiers in and plot them. This defines the region D in the ROC space. Further, we get the ROC convex hull region D of D' .

S3) We find the tangential point between the line with slope in (2) and the ROC convex hull region D'. The classifier corresponding to the found tangential point is the most cost-effective classifier.

## 8. COST MINIMIZATION IN HMM-BASED IMAGE SEGMENTATION

In this section, we find the HMM-based classifier to minimize the (expected) cost of misclassification (1) (ECM). To get better understanding of the problem, we begin with HMM without spatial coherence such as GMM [1]. The model assumes that the testing image is composed of many

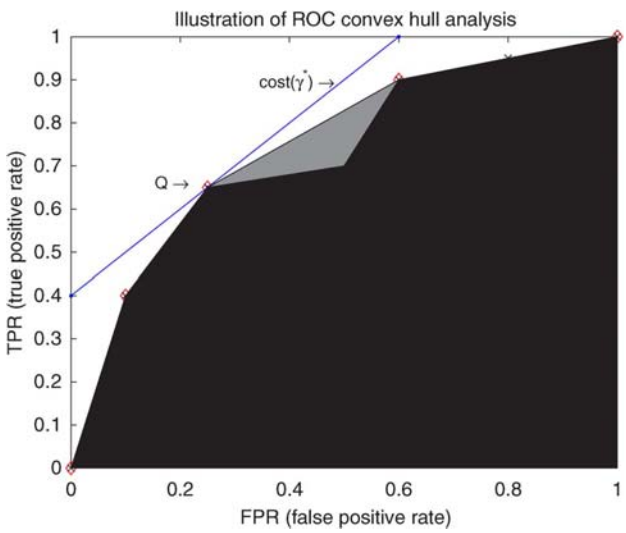


Fig. 4. ROC convex hull analysis: Given different misclassification cost, the intersection of ROC convex hull and iso-cost line provides the optimal classifier minimizing the total class-dependent misclassification cost.

independent sub-blocks, say $X_k$ s for k= 1,2, K . The model has unknown parameter γ which indicates the prevalence of each class. We let γ be the ratio of prior probability of positive class (class P) to that of negative class (class N). Then, given γ , the optimal classifier to minimize the ECM among all classifiers is the maximum a posteriori (MAP) classifier that assigns $x_k$ to class N, for each , if

$$F_n(x_k) / F_n(x_k) \geq (1/\alpha) \qquad (19)$$

where $F_n(X_k)$ and $F_n(X_k)$ is the probability density function of class N and class P, respectively. In GMM, both $F_n(X)$ and Fp(X) are density functions of the GMM. Suppose we denote the classifier in (3) given γ as C(γ) , and their collection as F . We further find that this collection is invariant to the cost α. In other words, we always have same collection D regardless of what α we choose.

In practice, to find the optimal classifier, we set α to be an arbitrary fixed number, and we consider the set of conditionally optimal classifiers which minimize the cost of misclassification given $\gamma \geq 0$ . We let D be the set of their FPRs and TPRs in the ROC space. Following the steps in the previous section, we get D'and find the tangent point $\gamma^*$ and between the line (2) and D'.

Now, we move to the HMM-based segmenter. The HMM based segmenter assumes that the hidden process is from the Markov random field having two unknown parameters γ and β; the parameter γ is the parameter that represents the magnitude of magnetization of the random field which implies the dominance of class P against class N in common words; it is also related to the ratio of the prior probability of P to N; β is the parameter to measure the strength of spatial coherence. For example, The HMGMM model uses the generalized bond percolation (BP) model. Let $Z=\{Z_{ij}, (i,j) \in i,j = 1,2,3,\ldots\ldots n\}$ with $Z_{ij=} -1$ or 1: $Z_{ij=} 1$ and $Z_{ij=} - 1$ implies the class P and N, respectively. The generalized BP assumes that the probability Z is

$$P(Z; \alpha,\beta)=1/\psi(\gamma,\beta) \exp \{ Zij \log \gamma + n_{con}(Z) \log\beta + n_{dis}(Z)\log(1-\beta) \qquad (20)$$

Where $n_{con}(Z)$ ($n_{dis}(z)$ ) is the number of concordant (dis-concordant) adjacent pairs which are neighbors to each other. Here, ψ (α,β) is the partition function that is

$$\sum_n \exp\{\sum Z_{ij} \log \gamma + n_{con}(Z) \log\beta + n_{dis} (Z) \log(1-\beta) \qquad (21)$$

As in the spatially uncorrelated model, we consider the collection F of MAP classifier C(γ,β) , which, given γ and β, assigns the observed image X= { $x_{ij}$, i ,j= 1,2,……,n}to Ž=arg max P(Z/x;γ,β). We let C(γ,β) be the MAP classifier, given γ and β, and F be the collection of C(γ,β)s. We further define D and D' similarly. To find the optimal classifier, again, we compute the tangential point between the line (2) and D', the convex hull of D in the ROC space.

In HMM-based segmenter, evaluation of D is computationally quite intensive in practice. Here, we introduce a suboptimal but fast algorithm for it. Many algorithms are studied from deterministic annealing to simulate annealing to Markov chain Monte Carlo method to find the MAP. They often assume γ=0 and estimate β , the parameter of spatial coherence in finding the MAP. For example, in [11], the parameter β is estimated using the maximum likelihood estimate (MLE) along with the Gibbs sampler. The MLE is $\beta\square^{MLE} = E\{n_{con}(Z|X)/$ total number of edges}, and we approximate the right hand side of the equation using Monte Carlo method to get the estimate.

We let β□(γ) be the estimate of β given γ, and we approximate the boundary points of D' with FPRs and TPRs of C(γ, β□(γ)) by moving . As stated before, knowing boundary dD is sufficient to find the tangential point between iso-cost lines and D. The approximated boundary, denoted by dD□ , is a curve from (0,0) to (1,1). Finally, dD□ is approximated by the convex hull of dD□, and the optimal classifier is found using the procedure in Section 2.

## 11. EXAMPLE

In this section, we apply the procedure in Section 5 tosegment the aerial image with HMM with generalized BW-GA model in 6. The aerial image is composed of many sub-blocks which are classified into "natural" regions or "man-made" regions. We call the "natural" sub-block as "negative" and "man-made" one as "positive." The HMMBWGA model has two parameters γ and β which reflects the overall portion of man-made and spatial coherence between adjacent sub-blocks, respectively. Each (γ, β) introduces a classifier, say C (γ ,β) ,and a point of (FPR, TPR) in the ROC space. We let be the collection of all these (FPR, TPR)s.

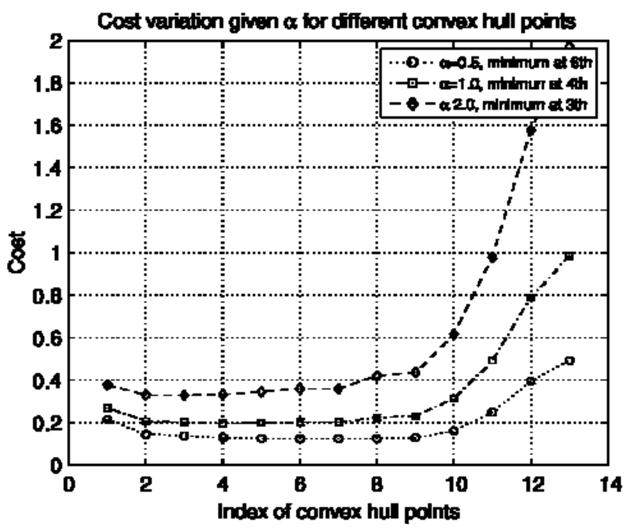


Fig.5 R O C convex hull analysis

In the experiment, we vary exp(γ)from 0.35 to 2.85 by 0.1, and we get the empirical FPRs and TPRs from the segmented results Z□ . We then compute the convex hull of (FPR, TPR)s of C(γ,β□(γ))s to approximate the convex hull of D. We denote the convex envelope as dD'. Fig. 4 plots C(γ, β(γ))s and their convex envelope in the ROC space.

We find the optimal classifier to minimize the misclassification cost (1) is the one that corresponds to the tangential point between the (2) and dD'□. The dD'□ is piecewise linear function with 11 supporting points (points that have changes in slope). Each point on is the optimal classifier to minimize the cost for a specific choice of α.

Now, we report some details of the analysis for ,α=0.5,1, and 2. First, Fig. 6 plots the misclassification cost of the 11 supporting points for each α = 0.5 , 1, or 2. It confirms that the chosen classifiers minimize the cost given α. Second, classifiers we consider, and be the set of their random mix-tures. Then, is the smallest convex region which contains . The goal of this letter is to find the classifier which minimizes the cost (1) among classifiers in . Since and given different misclassification b cost, the intersection of ROC convex hull and iso-cost line provides the optimal classifier minimizing time. the segmenter tends to classify both "natural" and "man-made" blocks to "man-made" ones. Finally, the original HMMBW-GA in [11] is the classifier with exp (γ) = 1 . Its FPR and TPR is 0.22 and 0.97, respectively, which is close to the (approximated) boundary of D' . Some computation shows that the HMMBW-GA is close to the optimal for between 0.16 and

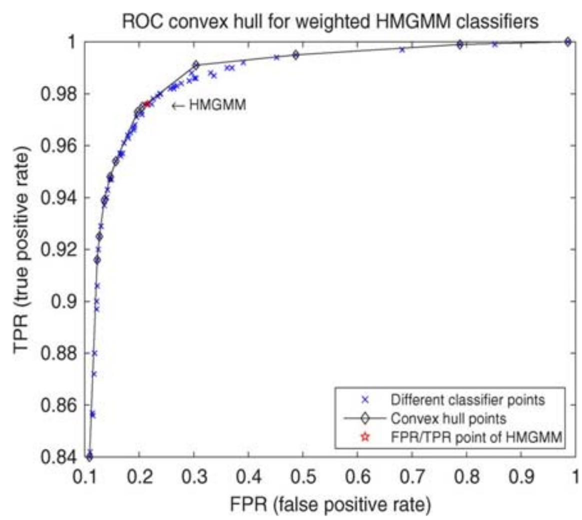


Fig 6: Roc convex curve : Plot FPR & TPR for all choices of fixed γ and its fixed β and its convex hull

The Baum-Welch algorithm gives you both the most likely hidden transition probabilities as well as the most likely set of emission probabilities given only the observed states of the model. The Baum-Welch algorithm is essentially the Expectation-Maximization algorithm applied to a HMM; as a strict EM-type algorithm you're guaranteed to converge to at least a local maximum, and so for unimodal problems find the MLE. It requires two passes over your data for each optimizes the Gaussian parameters considering as a fitness function the results of the classification application.

Results show the improvement of GA techniques for human activities recognition. algorithm that is based on the definition of the real observations as a mixture of two Gaussians for each state. The application of the GA follows the same principle but the optimization is carried out considering the classification. In this case, GA techniques for human activities recognition.

## 12. EXPERIMENTAL STUDY

The experimental study consist of using the HMM model parameter resulting from both B-W and hybrid GA-BW training systems, to classify a set of unknown image utterances. The image database used in the study comes with MixtGaussian toolbox download file, the training data contains digits from 1 to 15, each digit is repeated by 15 speakers, and the testing data contains the same digit repeated by 10 speakers. The recognition software is implemented in the MixtGaussian toolbox to. This test is repeated ten times as it is shown in table.

| Training technique | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| BW | 68 | 76 | 76 | 72 | 68 | 76 | 68 | 84 | 72 | 64 |
| GA-BW | 76 | 64 | 88 | 72 | 84 | 76 | 72 | 80 | 76 | 72 |

Table 1: Recognition rate for both BW and GA-BW within 10 experiments

From the results of table 1 it is difficult to statute about the quality of B-W and GA-BW algorithms in term of recognition rate. However it is clear from results from table2 that GA-BW algorithm performs better than the classical B-W algorithm. We believe that this improvement is due to global searching ability of GA . In the near future it will be used in continuous satellite aerial image classification.

| Training Technique | Minimum | Maximum | Average |
|---|---|---|---|
| BW | 64% | 84% | 72.67% |
| GA-BW | 64% | 88% | 76% |

Table 2: Minimum Maximum & Average of recognition rate for both BW & GA-BW

## CONCLUSION

The user can get a better image about the arial image and also tumour cells and location exactly and not randomly .A Hidden Markov Model (HMM) is used as an efficient and robust technique for human activities classification. The HMM evaluates a set of image recordings to classify each scene as a function of the future, actual and previous scenes. The probabilities of transition between states of the HMM and the observation model should be adjusted in order to obtain a correct classification. In this work, these matrixes are estimated using the well known Baum-Welch algorithm that is based on the definition of the real observations as a mixture of two Gaussians for each state. However, the B-W algorithm uses an initial random guess of the parameters, therefore after convergence the output tends to be close to this initial value of the algorithm, which is not necessarily the global optimum of the model parameters. In this project, a Adaptive Baum-Welch (GA-BW) is proposed. The application of the GA follows the same principle but the optimization is carried out considering the classification. In this case, GA techniques for image activities recognition. Under the same conditions the hybrid GA-BW algorithm performs better than Baum-Welch method. We believe that this improvement is due to global searching ability of GA . In the near future it will be used in continuous satellite aerial image classification.

## Acknowledgment

The authors would like to thank their parents and relations for the financial support received .They would like to thank the reviewers and editors of IJCSI. Finally, they appreciate all the readers who spend their precious time to read this paper.

## REFERENCES

[1] K. Won, A. Prugel-Bennett, A. Krogh, "Evolving the Structure of Hidden Markov Models," IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION, VOL. 10, NO. I, FEBRUARY 2008.

[2] K. Pyun, J. Lim, C. S. Won, and R. M. Gray, "Image segmentation using hidden Markov Gauss mixture models," *IEEE Trans. Image Process.*, vol. 16, no. 7, pp. 1902–1911, Jul. 2007.

[3] Johan Lim and Kyungsuk (Peter)Pyun, "Cost-Effective Hidden Markov Model-Based Image Segmentation", VOL. 16, NO. 3, MARCH 2009.

[4] S. Kwong, C.W. Chau, K.F. Man, K. S. Tang, "Optimization of HMM topology and its model parameters by genetic algorithms," Pattern Recognition Society, V. 34, pp.S09-S22,2009.

[5] K. Won, A. Prugel-Bennett, A. Krogh, "Training hmm structure with genetic algorithm for biological sequence analysis,", Bioinformatics 20(1 8),3613-3619,2009.

[6] L. Liu, H. Huo, B. Wang chev, P. Woodland, The HTK Book (for HTK Version 34.1), Cambridge University Engineering Department, 2010

**R.RAJASREE** received the Bachelor of Engineering in Electronics and Communication Systems from Anna University Chennai 2009 and now she doing her Master of Engineering Communication Systems in Anna University Tirunelveli, She is a student of PSN College of Engineering and Technology. She has been participated many national conferences and seminars and she presented 5 papers in national conferences.Her research interests include Digital Image Processing and Image Segmentation as well as in Networking.

**J.NALINI** received the M. Sc Electronics , M. Phill Electronics Manonmaniyam University 2006 , M.E Communication Systems from Vinayaka Mission University 2008 . She is an Assistant professor of Department of Electronics and Communication Systems in the PSN College of Engineering and Technology. Her research interests include Digital Image Processing, and wide area networking.

**S.C.Ramesh** is an Associate professor of Department of Aeronautical engineering. He has been served on program committees and organizing committees of various international conferences and symposia. He is a member of IEEE.