

# Hidden Markov Models Applied to Snakes Behavior Identification

Wesley Nunes Gonçalves, Jonathan de Andrade Silva, Bruno Brandoli  
Machado, Hemerson Pistori, and Albert Schiaveto de Souza

Dom Bosco Catholic University  
Research Group in Engineering and Computing  
Av. Tamandaré, 6000, Jardim Seminário, 79117-900, Campo Grande, MS, Brazil  
{wnunes, jsilva, bmachado}@acad.ucdb.br  
{pistori, albert}@ucdb.br  
<http://www.gpec.ucdb.br>

**Abstract.** This paper presents an application of the hidden Markov models (HMMs) to the recognition of snakes behaviors, an important and hard problem that, as far as the authors know, has not been tackled before, by the computer vision community. Experiments were conducted using different HMM configurations, including modifications on the number of internal states and the initialization procedures. The best results have shown a 84% correct classification rate, using HMMs with 4 states and an initialization procedure based on the K-Means algorithm.

**Key words:** Hidden Markov Models, Animals Behavior Recognition, Snakes

## 1 Introduction

Snakes were suggested as one of the main groups of animals for the evaluation of ecological and evolutionary hypotheses [1]. Moreover, this group is an interesting and important producer of venom, which are intensively used in the development of drugs for hypertension control, analgesic, anticoagulants, among others. So, research on snakes behavior, in order to understand its impact on venom production in natural and controlled habitats, has been considerably increased in the last years.

The habitat for reptiles, in particular snakes, is a much unexplored and little understood topic. The creation of snakes in captivity is a difficult task that can benefit from the analysis and identification of snakes behaviors in natural and artificial environments. The behavior of the snakes is influenced by many factors, like temperature, solar radiation, humidity and cycle of seasons. The ability to identify and predict snakes activity is essential in the process of venom production.

The identification of animals behavior is usually carried out through a non-automatic procedure involving high time consuming sections of visual observation and annotations. The reproducibility and precision of this procedure is

affected by the fatigue and distraction of the observers. The automation of this process, using computer vision techniques, could bring to the area much more reliable results and the possibility to extend the experiments to situations and environments that are not easily accessible using the current methods. The benefits are even higher for snakes, which presents some activities that, to be observed, would require many hours of continuous observation. Automation could also provide information, related to precise physical measurements, that can not be achieved by naked eyes observation [2].

This work evaluated the use of hidden Markov models to automatically identify a particular snake behavior that occurs during an attack. The species of snakes were used during the experiments: Common boa (*Boa constrictor*), Neotropical Rattlesnake (*Crotalus durissus terrificus*) and Pitviper (*Bothrops jararaca*). The HMM model was trained using 1000 frames from 20 image shots of the snakes during the attack and not attacking. The model was tested on 10 image shots not presented in the training data. A semi-automatic procedure, based on SVM (Support Vector Machines) supervised learning [3], was adopted, in the segmentation phase, to separate the snake region from the background. Parameter extraction were based on Image Moments. The best correct classification and execution time was achieved through the variation of some of the parameters of the hidden Markov models, like the number of states and the number of iterations used by the Baum Welch parameters estimation procedure. Moreover, two different initialization techniques were explored during the HMM's parameter estimation. The attack behavior was inferred with accuracy of 84% by a HMM with 4 internal states.

The paper is organized as follows. Section 2 presents related work that applies the hidden Markov models to computer vision problems. Section 3 briefly reviews the image moments parameter extractor. In Section 4 the hidden Markov Models used in this work is explained. The experiments performed and the results are shown in details in Sections 5 e 6, respectively. Finally, the conclusion and future work are discussed.

## 2 Previous Work

The hidden Markov models (HMMs) have been used in many areas, mainly in systems for speech recognition [4], behaviors recognition [5] and hand-write recognition [6]. In [7] HMM is applied to 2D objects recognition in images. The HMM, jointly with the contour invariant feature, have been tested in four different objects. For each object, a HMM was estimated using a set of fifty training images. The classification was carried out using ten images for each object, resulting in a correct classification rate of 75%.

Starner and Pentland [8] describe a system for recognition of American sign language using HMM. The correct classification rate was 99.2% for words, however, the feature set proved to be limited and the gestures were expect to occur in pre-specified spatial positions in the image, as hands positions were not normalized. A new technique for character recognition is presented in [9]. The features

are extracted from a gray level image and a HMM is modeled for each character. During the recognition, the most probable combination of models is found for each word, using a technique based on dynamic programming.

In [10] a HMM for recognition of faces is described. The image containing the face is divided in five blocks (hair, forehead, eyes, nose and mouth) and each block is represented by a HMM internal state. The feature vectors are obtained from each block by the Karhunen-Loeve transform approach. The HMMs are frequently used to describe a sequence of patterns characterizing a behavior. In [5], human behaviors are identified. Those behaviors are related to legal and illegal activities, captured by a camera, and carried out in an archaeological site. For the identification of those behaviors, the images are segmented utilizing movement detection and followed by a shade removal processing. After that, the human posture is identified using histogram and similarity measure based in the Manhattan distance. Behavior recognition is carried out by HMMs, where the internal states represents different postures. Experiments were carried out in the identification of four behaviors and a mean correct classification rate of 86,87% was reported.

In [11] an animal behaviors classification system is presented. That system uses a combination of HMM and kNN for learning to recognize some movements. The system was evaluated in bees paths extracted from a 15 minutes image sequence. A classification rate of 81.5% has been reported on this problem.

### 3 Image Moments

An image can be modelled as a 2D discrete function  $I$ , where the intensity of each pixel is indexed as  $I(x, y)$ . Equation 1 represents the image regular moments of order  $p, q$  of an image.

$$M_{pq} = \sum_{x=1}^{width} \sum_{y=1}^{height} x^p y^q I(x, y) \quad (1)$$

Regular image moments can be used to represent some important properties of an object presented in an image, like the object area,  $M_{00}$ , and its center of mass,  $(\frac{M_{10}}{M_{00}}, \frac{M_{01}}{M_{00}})$ .

$$u_{pq} = \sum_{x=1}^{width} \sum_{y=1}^{height} (x - \bar{x})^p (y - \bar{y})^q I(x, y) \quad (2)$$

Central image moments, as defined in equation 2, can also be used to calculate some other interesting object properties, like its variance in X and Y axes (Equation 3), direction (Equation 4) and eccentricity (Equation 5).

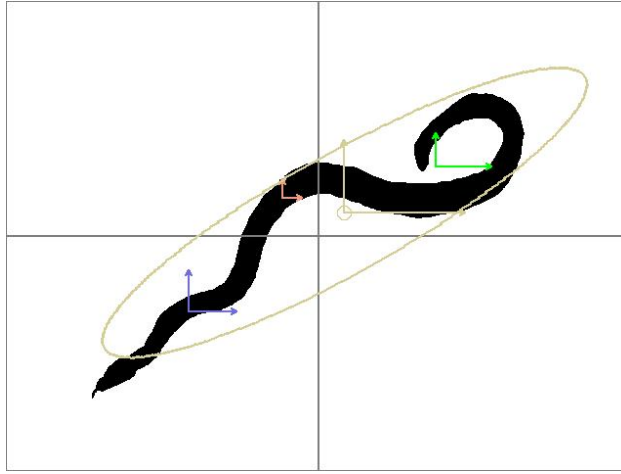
$$\sigma_x^2 = \frac{u_{20}}{m_{00}}, \quad \sigma_y^2 = \frac{u_{02}}{m_{00}} \quad (3)$$

$$\theta = \frac{u_{02} - u_{20} - 2u_{11} + \lambda}{u_{02} - u_{20} + 2u_{11} - \lambda} \quad (4)$$

$$e^2 = \frac{u_{20} + u_{02} + \lambda}{u_{20} + u_{02} - \lambda} \quad (5)$$

$$\lambda = \sqrt{(u_{20} - u_{02})^2 + 4u_{11}^2} \quad (6)$$

Besides each of the above mentioned properties, calculated for the whole object, in this work, following a methodology suggested by Freeman [12], the object is further divided in 4 equal regions, and for each of these regions, the same image moments properties are calculated. In this way, the system can combine global and local information during the classification phase. Figure 1 illustrates, using an image moments visualization tool, the properties extracted from a snake image that has been previously segmented and binarized.



**Fig. 1.** Images Moments Application.

## 4 Hidden Markov Models

The hidden Markov models (HMMs) are used to model a pair of complementary stochastic processes. The first process is represented by a set of unobserved states, also called hidden or internal states. In the special case of first order HMMs, the current state of the system depends only on the previous state, and the probability distribution that models states transition is usually represented via a transition matrix  $A = \{a_{ij}\}$ , with

$$a_{ij} = P(q_{t+1} = S_j | q_t = S_i) \quad 1 \leq i, j \leq N \quad (7)$$

where  $N$  is the number of states,  $q_t$  is the current state of the system and  $S = \{S_1, S_2, S_3, \dots, S_N\}$  is the set of hidden states.

The second stochastic process models the probability of observing or measuring some predetermined values (the observed values or symbols) given that the system is in a specific state (hidden). A sequence of  $T$  observations is represented by a set  $O = O_1, O_2, O_3, \dots, O_T$ , where each element  $O_t$  is a member of the symbols set  $V = v_1, v_2, \dots, v_M$ . The emission or observation probability of any symbol given an internal state  $j$  is defined by a matrix  $B = \{b_j(k)\}$ , with

$$b_j(k) = P(O_t = v_k | q_t = S_j) \quad 1 \leq j \leq N, \quad 1 \leq k \leq M \quad (8)$$

The initial probability of each state is represented by a set  $\pi = \{\pi_i\}$ , with

$$\pi_i = P(q_1 = S_i) \quad 1 \leq i \leq N, \quad \text{with} \quad \sum_{i=1}^N \pi_i = 1 \quad (9)$$

The problem of estimating the parameters (probability matrices) of a HMM is usually called learning problem, whereas the problem of calculating the likelihood of an observation sequence given a particular HMM is called evaluation problem. Given an observed sequence  $O$  and a HMM model  $\lambda = (A, B, \pi)$ , the evaluation problem is to calculate  $P(O|\lambda)$ . One of the procedures that resolves this problem efficiently, based on dynamic programming, is known as the Forward-Backward algorithm. This procedure defines a variable  $\alpha_t(j)$  (the forward variable) that represents the probability of a partial observation sequence (from time 0 to  $t$ ) given an state  $S_j$  (in time  $t$ ) and the model  $\lambda$ . The variable is updated incrementally using the recursive procedure defined by Equations 10 and 11 until the full observation sequence is reached and  $P(O|\lambda)$  can be easily calculated using Equation 12.

$$\alpha_1(j) = \pi_j b_j(O_1), \quad 1 \leq j \leq N. \quad (10)$$

$$\alpha_{t+1}(j) = \left[ \sum_{i=1}^N \alpha_t(i) a_{ij} \right] b_j(O_{t+1}), \quad 1 \leq t \leq T - 1. \quad (11)$$

$$P(O|\lambda) = \sum_{j=1}^N \alpha_T(j) \quad (12)$$

In order to calculate the backward variable  $\beta_t(i)$ , representing the probability of partial observations from  $t + 1$  until  $T$  given the state  $S_i$  in the time  $t$  and a model  $\lambda$ , a similar procedure is followed, but in a reverse manner. The procedure is summarized in equations 13, 14 and 15.

$$\beta_T(i) = 1, \quad 1 \leq i \leq N. \quad (13)$$

$$\beta_t(i) = \sum_{j=1}^N a_{ij} b_j(O_{t+1}) \beta_{t+1}(j), t = T-1, T-2, T-3, \dots, 1 \quad 1 \leq i \leq N. \quad (14)$$

$$P(O|\lambda) = \sum_{i=1}^N \pi_i b_i(O_1) \beta_1(i) \quad (15)$$

In spite of the somewhat misleading name, in order to solve the evaluation problem, one must choose to use the forward or the backward procedure, not both. The learning problem is to choose the parameters of the model  $\lambda = (A, B, \pi)$ , that maximizes locally  $P(O|\lambda)$ . A well-known algorithm that solves the problem in polynomial time is the Baum-Welch, a specialization of the EM algorithm. Detailed information on the Baum-Welch algorithm and HMMs in general can be found in [4, 13, 8, 9].

## 5 Experiments and Results

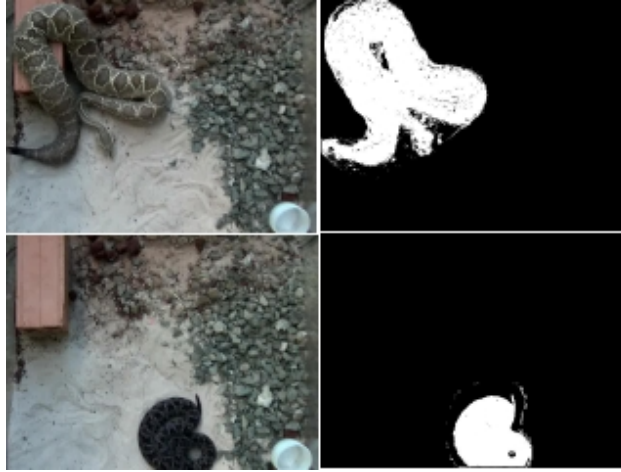
The experiments with the HMMs were carried out with 30 images shot representing the presence and the absence of the attack behavior. An example of the attack behavior can be visualized in Figure 2. The images were captured using a TRENDNET TV-IP301W camera with a spatial resolution of 640 x 480 pixels. The snake was held in a place that simulates its natural environment and the pictures were taken from above.



**Fig. 2.** Four frames of an attack sequence

Snakes use camouflage to hide from enemies and to more easily capture other animals, which turns the segmentation problem, in this context, very difficult.

A supervised learning strategy, based on Support Vectors Machine (SVM), were used to separate the snake from the background. For each image, a rectangular region surrounding the snake was manually determined in order to turn the segmentation phase easier. Color based attributes, extracted from snake and background regions, were used to feed the learning process. The results of this segmentation procedure is illustrated in Figure 3.



**Fig. 3.** Image Segmentation for Snakes.

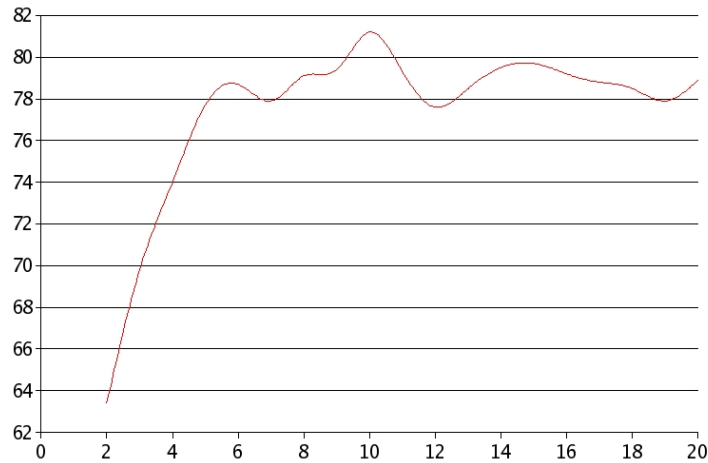
After segmentation, image moments based attributes were used to extract information related to the shape of the snake in each frame. These attributes were further discretized using the vector quantization LBG algorithm [14] and used as observation symbols for two HMMs, one corresponding to attack behavior and the other to non-attack. A total of 20 images sequences were used for training and 10 for testing the classification module based on HMM. All the experiments were carried out in a computer with a P4 2.8GHz processor, 512MB of RAM and Fedora Core 5.

Experiments were conducted in order to find the HMM configuration that gives the higher correct classification rate in the problem of attack and non-attack behavior classification. Three parameters were evaluated: the number of hidden states, the number of iterative steps of the Baum-Welch algorithm and the initialization procedure during the learning phase. The number of hidden states and iterative steps varied from 2 to 20, and from 100 to 1000, respectively. Two initialization procedures were tested: the procedure suggested in [4], that assumes a uniform probability distribution for all matrices and a K-Means [15] based approach. For the K-Means approach, the  $A$  e  $\pi$  matrices are calculated as in [4], however, the  $B$  matrix is initialized in a different way. First, the training set is clustered using K-means with  $K$  being the number of internal states. Then,

using the mapping from states to training samples generated by the clustering procedure, the  $B$  matrix can be estimated by simple counting techniques.

The results are presented in Figures 4, 5, 6 and 7. The correct classification rate in the graphic are calculated through on average of the results of HMM configurations. Like this, the correct classification rate for the a HMM with 20 states is on average of the result of all the HMM configurations with 20 states.

In Figure 4 the graphic relates the number of states to the correct classification rate. The best rate, of 81.2%, has been reached at 10. This performance has been achieved at a 515.41ms execution time cost (learning and evaluation time), as the graphic in Figure 5 indicates. The graphic in Figure 5 also illustrates the polynomial time complexity, on the number of states, of the learning and evaluation algorithms for HMMs. Nonetheless, for real time application, like, for instance, in a system that should immediately react to a snake attack, even a polynomial performance could not be sufficient, and a balance among accuracy and time response should be persued. In all graphics, the values of the parameters not shown, including the initialization approach, were chosen as the ones that give the mean performance.

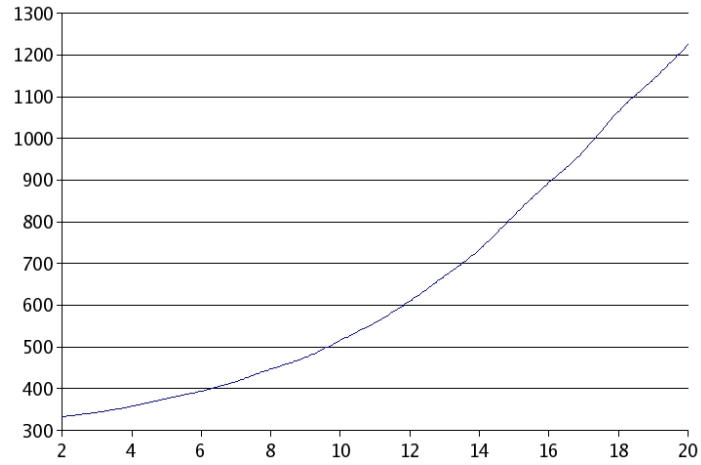


**Fig. 4.** Number of States X Correct Classification Rate

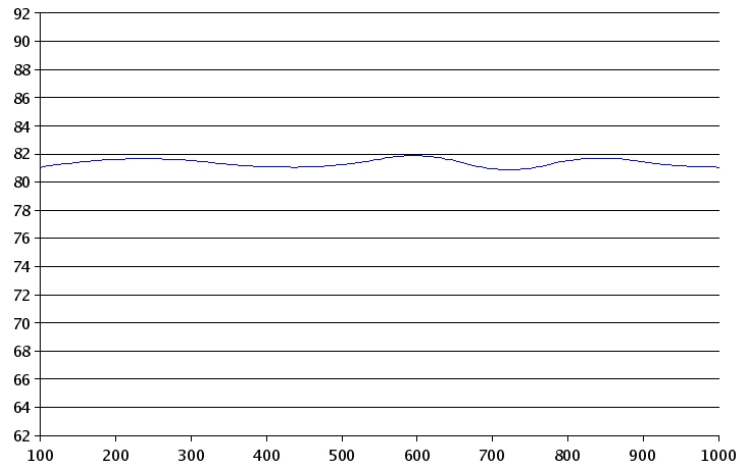
The graphic of Figure 6, relating the number of iterations and the correct classification rates, indicates that this parameter does not have a great effect on the performance, that is always kept near the 81% value.

Finally, in Figure 7 a graphical comparison between the two initialization approaches is presented. The graphic shows the relation on the number of states and correct classification rate for each approach. The standard initialization approach (uniform distribution assumption) outperforms the K-Means based approach when the number of states is greater than 11. Below this value, the



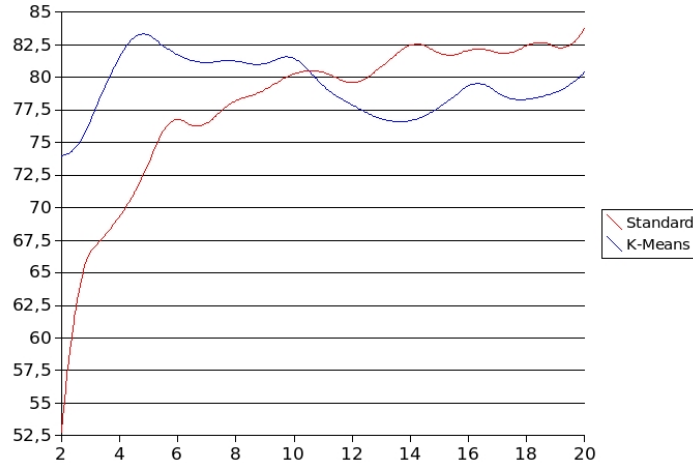


**Fig. 5.** Number of States X Execution Time (in milliseconds)



**Fig. 6.** Number of Iterations X Correct Classification Rate

K-Means approach is always better. This is due to the fact that the clustering performed before the initialization reviews to the learning module that some hidden states are not associated (or highly associated) to any observation. The best correct classification rate, 84%, of all experiments, was achieved using the K-Means initialization approach, with only 4 states.



**Fig. 7.** Number of States X Correct Classification for the Two Different Initialization Approaches

## 6 Conclusion and Future Works

This paper showed an application of hidden Markov models to the recognition of snakes behavior. Experiments in the number of internal states, Baum-Welch iterations and initializations approaches were conducted in order to find the best configuration for this particular problem. A maximum correct classification rate of 84% has been achieved.

For future research, it would be interesting to include information related to the contour of the snake in the feature vectors, and to evaluate other HMM types, like the ones that use continuous probability distribution to overcome that problem of having to discretize the values observed. It is also important to expand the tests using a larger amount of images, with different kinds of animals and environments.

**Acknowledgments.** This work has received financial support from Dom Bosco Catholic University, UCDB, Agency for Studies and Projects Financing, FINEP, and Foundation for the Support and Development of Education, Science and

Technology from the State of Mato Grosso do Sul, FUNDECT. One of the co-authors holds a Productivity Scholarship in Technological Development and Innovation from CPNQ, the Brazilian National Counsel of Technological and Scientific Development, and some of the other co-authors have received PIBIC/CNPQ scholarships.

## References

1. Rivas, J.A., Burghardt, G.M.: Snake mating systems, behavior, and evolution: The revisionary implications of recent findings. *Journal of Comparative Psychology* **119**(4) (2005) 447–454
2. Spink, A.J., Tegelenbosch, R.A.J., Buma, M.O.S., Noldus, L.P.J.J.: The ethovision video tracking system—a tool for behavioral phenotyping of transgenic mice. *Physiology and Behavior* **73**(5) (2001) 731–744
3. Burges, C.J.C.: A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery* **2**(2) (1998) 121–167
4. Rabiner, L.R.: A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE* **77** (1990) 257–286
5. Leo, M., D’Orazio, T., Spagnolo, P.: *Human Activity Recognition for Automatic Visual Surveillance of Wide Areas*. 1 edn. Academic Press (1999)
6. Hu, J., Brown, M.K., Turin, W.: Hmm based on-line handwriting recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **18**(10) (1996) 1039–1045
7. Hornegger, J., Niemann, H., Paulus, D., Schlottke, G.: Object recognition using Hidden Markov Models. In Gelsema, E.S., Kanal, L.N., eds.: *Pattern Recognition in Practice IV: Multiple Paradigms, Comparative Studies and Hybrid Systems*. Volume 16., Amsterdam, Elsevier (1994) 37–44
8. Starner, T., Pentland, A.: Visual recognition of american sign language using hidden markov models. Technical Report Master’s Thesis, MIT, Program in Media Arts & Sciences, Massachusetts Institute of Technology, Cambridge, USA (1995)
9. Aas, K., Line, E., Tove, A.: Text recognition from grey level images using hidden Markov models. *Lecture Notes in Computer Science* **970** (1995) 503–508
10. Nefian, A.V., Hayes, M.H.: Face detection and recognition using hidden markov models. In: *ICIP* (1). (1998) 141–145
11. Feldman, A., Balch, T.: Automatic identification of bee movement. Technical report, Georgia Institute of Technology, Atlanta, Georgia 30332, USA (2003)
12. Freeman, W., Tanaka, K., Ohta, J.: Computer vision for computer games. In: *Int’l Workshop on Automatic Face- and Gesture-Recognition*, Killington, Vermont, USA (1996)
13. Montero, J.A., Sucar, L.E.: Feature selection for visual gesture recognition using hidden markov models. In: *ENC*. (2004) 196–203
14. Shen, F., Hasegawa, O.: An adaptive incremental lbg for vector quantization. *Neural Netw.* **19**(5) (2006) 694–704
15. Malyszko, D., Wierzchon, S.T.: Standard and genetic k-means clustering techniques in image segmentation. *cisim* **0** (2007) 299–304