

Unsupervised Segmentation of Naval Infrared Images through a Markov Random Field Model

Dietmar Wippig, Bernd Klauer and Hans Christoph Zeidler

Department of Electrical Engineering

Helmut-Schmidt-University

22043 Hamburg, Germany

{dietmar.wippig, bernd.klauer, h.ch.zeidler}@hsu-hamburg.de

Abstract—Ship recognition systems use infrared imaging technology at night or at poor visibility. On the infrared images taken in offshore environments three kinds of regions can be distinguished: ship, sky and sea. For the classification of ships the area in the image covered by the ship contour has to be extracted. The purpose of image segmentation is to partition an image into constituent components for further processing such as recognition. There are two main categories of segmentation algorithms: supervised and unsupervised. Supervised segmentation can be used where the features are known a priori. These algorithms classify data inputs based on training data sets or on user inputs, while unsupervised segmentation is only based on the features in the current data set. Because the features are unknown and the procedure should be independent from user interaction to get reproducible results, unsupervised segmentation is chosen. A very effective technique for unsupervised segmentation is clustering. Therefore this paper investigates common clustering algorithms for their suitability. The Markov Random Field Model is found to deliver the best results for segmenting infrared images but also has the highest computational demand.

Keywords: Markov Random Field, Image Segmentation, Infrared Images

1. Introduction

Infrared imaging technology supports military and civil monitoring of navigation and ship recognition systems. Ship recognition systems use the ship contour for classification which has to be extracted from the infrared image. Mostly this task is performed in offshore environments. Due to their different temperature three kinds of regions occur in the infrared images: ship, sky and sea. The sea can be warmer than the sky and vice versa. That does not matter, because the ship is always the warmest region. Depending on the polarity the ship is the brightest or the darkest region in the image (Fig. 1).

Former approaches allowing to perform a ship focused segmentation of infrared images are mainly based on histogram thresholding [1] [2].

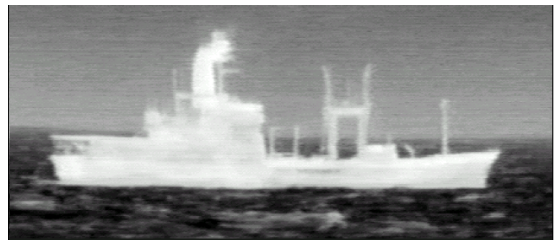


Fig. 1. Example of a infrared ship image

These approaches were not applicable to our images, because of the poor image quality obtained from the IR imaging devices. Other concepts investigated at the Helmut-Schmidt-University were based on a statistical approach [3] and on region growing [4], but using image sequences to identify the ship region. Because results should be independent from user interaction modern supervised image segmentation algorithms like graph cut [5] cannot be used which can be appropriately adapted. Therefore this paper focuses on unsupervised segmentation by clustering algorithms. The clustering performance of K-means, Fuzzy C-means, Gaussian Mixture Models and Markov Random Field Models is investigated.

2. Clustering

Clustering is not only applicable in image segmentation. It is a common approach for grouping data. Clustering algorithms try to organize feature data into clusters using a given measure of similarity. The feature data within a cluster should be more similar to each other than to feature data belonging to other clusters. The distribution of the feature data and the number of clusters are often unknown and therefore they have to be estimated. The similarity measure can also only be an assumption. Therefore there is no single solution for clustering algorithms.

2.1. K-means Clustering

K-means [6] is one of the simplest unsupervised data clustering algorithms. Given a certain number of clusters (K) the feature data are grouped to K clusters based on the cluster's centroid. The algorithm attempts to find centroids where the sum of the distances to the clusters feature data will be a minimum.

The K-means algorithm works as follows:

1. Randomly choose K feature data as initial centroids.
2. Assign each feature value to the cluster that has the closest centroid.
3. Calculate the new cluster centroids.
4. Repeat steps 2 and 3 until the centroids no longer move.

A common approach to find the distance between feature data is to use the Euclidean distance. Another distance model that can be used is the Mahalanobis distance using the variance of the cluster. The Mahalanobis distance from values with means $\mu = \{\mu_1, \dots, \mu_n\}$ and Covariance matrix Σ for $x = \{x_1, \dots, x_n\}$ is given by

$$D_M = \sqrt{(x - \mu)^T \Sigma^{-1} (x - \mu)}.$$

2.2. Fuzzy C-means Clustering

Another clustering algorithm is the Fuzzy C-means (FCM) clustering [7] [8] based on the fuzzy set theory by Zadeh [9]. In the FCM approach, feature data can belong to more than one cluster. The degree of membership of each feature data to such a unsharp cluster can be determined. The algorithm tries to minimize the total distance between feature data and cluster means weighted by the corresponding membership values. For each point i C fuzzy membership numbers u_{ij} are derived. They represent the degree of membership of i th feature data to k th cluster with $1 \leq i \leq N$ and $1 \leq j \leq C$, where N is the number of feature data and C is the number of clusters.

The following is a description of the Fuzzy-C-means algorithm, which can be found in [10].

1. Initialise the membership matrix $U^{(0)} = [u_{ij}] \forall i, \forall j$ and set the iteration

index t to 0.

2. At step t calculate the new cluster center vectors $v_j^{(t)} = [v_j]$ with the membership matrix $U^{(t)}$ by

$$v_j = \frac{\sum_{i=1}^N (u_{ij})^m \cdot x_i}{\sum_{i=1}^N (u_{ij})^m},$$

where $m \in [1, \infty)$ is the membership weighting exponent.

3. Calculate the updated fuzzy membership matrix $U^{(t+1)}$ by using $U^{(t)}$ and $v_j^{(t)}$

$$u_{ij} = \frac{1}{\sum_{k=1}^C \left(\frac{\|x_i - v_j\|}{\|x_i - v_k\|} \right)^{\frac{2}{m-1}}},$$

where $\|\cdot\|$ is any product norm.

4. If $\|U^{(t+1)} - U^{(t)}\| = \max_{ij} |u_{ij}^{(t+1)} - u_{ij}^{(t)}| < \epsilon$, then stop, otherwise set $t = t + 1$ and continue with step 2. ϵ is a termination criterion between 0 and 1.

Like in the K-means algorithm the Mahalanobis distance can be used instead of the Euclidean distance. This leads to the Gustafson-Kessel algorithm [11], where additionally the fuzzy covariance matrix is calculated after calculating the cluster center vectors. The fuzzy covariance for each cluster j is defined by

$$F_j = \frac{\sum_{i=1}^N (u_{ij})^m (x_i - v_j)(x_i - v_j)^T}{\sum_{i=1}^N (u_{ij})^m}.$$

2.3. Gaussian Mixture Model

A Gaussian Mixture Model (GMM) is the weighted sum of different Gaussian distributions representing different domains in feature space. The probability density function of an observation x given M component densities is

$$p(x|\Theta) = \sum_{j=1}^M \alpha_j \cdot p_j(x|\theta_j),$$

where $\Theta = (\alpha_1, \dots, \alpha_M, \theta_1, \dots, \theta_M)$ contains the parameters of the mixture PDF. α_i are the mixing coefficients and θ_i are the parameters of the density functions p_i . Each mixture component is d -dimensional Gaussian distributed with mean

μ_i and covariance matrix Σ_i :

$$p_i(x, \mu_i, \Sigma_i) = \frac{1}{|\Sigma_i|^{1/2} (2\pi)^{d/2}} \cdot e^{-\frac{1}{2}(x-\mu_i)^T \Sigma_i^{-1} (x-\mu_i)}$$

In clustering, GMMs can be used to estimate the relative contribution of each domain (or cluster) to the overall probability density function, so that the mean of each Gaussian distribution is an estimate of the cluster centroid. Therefore a clustering algorithm has to find the parameters of the GMM and to determine the membership of each feature data to a Gaussian distribution. Gaussian Mixture Models are often estimated using the Expectation-Maximization (EM) algorithm [12].

The EM algorithm is a general method to find the maximum likelihood estimation of the parameters of an underlying distribution from a given data set treated incomplete. Together with hidden variables and parameters associated with the problem the complete data are constituted. The EM algorithm consists of two steps to maximize the likelihood of the complete data. The expectation step (E-step) uses the current estimate of the parameters to compute the expected value of the complete-data log-likelihood. Then the maximization step (M-step) estimates new parameters to maximize the expectation. So each iteration increases the log-likelihood. The estimation steps are repeated until a local maximum is found.

The EM algorithm applied to GMMs can be summarized as follows [13]:

1. Initialise the parameters for the mixture of Gaussians and set the iteration index t to 0.
2. E-Step: calculate for all x_1, \dots, x_N and $l = 1, \dots, M$

$$\begin{aligned} p(l|x_i, \Theta^{(t)}) &= \frac{p(x_i|l, \Theta^{(t)})}{\sum_{j=1}^M p(x_i|j, \Theta^{(t)})} \\ &= \frac{\alpha_l^{(t)} \cdot p(x_i, \mu_l^{(t)}, \Sigma_l^{(t)})}{\sum_{j=1}^M \alpha_j^{(t)} \cdot p(x_i, \mu_j^{(t)}, \Sigma_j^{(t)})} \end{aligned}$$

where

$$p(x_i, \mu_l^{(t)}, \Sigma_l^{(t)}) = \frac{1}{|\Sigma_l^{(t)}|^{1/2} (2\pi)^{d/2}} \cdot e^{-\frac{1}{2}(x-\mu_l^{(t)})^T \Sigma_l^{(t)-1} (x-\mu_l^{(t)})}$$

3. M-Step: estimate the parameters

$$\alpha_l^{(t+1)} = \frac{1}{N} \cdot \sum_{i=1}^N p(l|x_i, \Theta^{(t)})$$

$$\left(\mu_l^{(t+1)}\right)^2 = \frac{\sum_{i=1}^N x_i \cdot p(l|x_i, \Theta^{(t)})}{\sum_{i=1}^N p(l|x_i, \Theta^{(t)})}$$

$$\left(\Sigma_l^{(t+1)}\right)^2 = \frac{\sum_{i=1}^N p(l|x_i, \Theta^{(t)}) (x_i - \mu_l^{(t+1)}) (x_i - \mu_l^{(t+1)})^T}{\sum_{i=1}^N p(l|x_i, \Theta^{(t)})}$$

4. If $\max_l |\alpha_l^{(t+1)} - \alpha_l^{(t)}| < \epsilon$ and $\max_l |\mu_l^{(t+1)} - \mu_l^{(t)}| < \epsilon$, then stop, otherwise set $t = t + 1$ and goto step 2. ϵ is a termination criterion between 0 and 1.

Because the EM algorithm only guarantees a local maximum a proper initialization is necessary. Often the result of K-means algorithm is taken to initialise the EM algorithm.

2.4. Markov random field Model

A Markov random field (MRF) model can improve previously described clustering approaches by considering the labels of the neighbors. Geman and Geman [14] adopt this approach to image processing. There are numerous MRF models applied to unsupervised image segmentation like [15] [16].

The key idea is to model the image pixels with their values and their presence and orientation of edges as atoms or molecules in a lattice physical system whose energy function can be determined. Because this energy function is Gibbs distributed, the image model is a MRF.

The image model based on considering the image segmentation problem in a Bayesian framework:

$$P(F = f|G = g) = \frac{P(F=f)P(G=g|F=f)}{P(G=g)},$$

where F is the segmented and G is the observed image.

Only the a priori probability $P(F = f)$ and the conditional observation probability $P(G = g|F = f)$ have to be calculated because $P(G = g)$ is constant and only the relative probability is required for maximizing the a posteriori probability $P(F = f|G = g)$. This *maximum a posteriori* (MAP) estimation is often maximized using the logarithm form

$$\ln P(F|G) \propto \ln P(F) + \ln P(G|F).$$

The negative logarithmic terms are treated as energies:

$$E = E_R + \alpha E_F,$$

where α is a weighting parameter that controls the effect of prior energy E_R and observation energy E_F to the entire energy E . Thus the maximization problem becomes a minimization problem.

For the prior (label) distribution energy the second order pairwise MLL (multi-level logistic) model is often taken. The energy of the pairwise MLL model is given by:

$$E_R = \sum_{i,j} \left(\beta \sum_{(k,l) \in N_{i,j}} \delta(f_{i,j}, f_{k,l}) \right),$$

where

$$\delta(f_{i,j}, f_{k,l}) = \begin{cases} -1 & , f_{i,j} = f_{k,l} \\ 1 & , f_{i,j} \neq f_{k,l} \end{cases},$$

$N_{i,j}$ is a centered neighborhood window around each image pixel $g_{i,j}$, $f_{i,j}$ is the label at position i, j and β is a positive constant.

The observation density is modeled as a Gaussian distribution with different means μ_m^k and standard deviations σ_m^k . So the means μ_m^k and standard deviations σ_m^k of each region m in the k th feature component are also slowly varying functions of i, j . The corresponding energy can be formulated as

$$E_F = \sum_{i,j,m=F_{i,j}} \left\{ \sum_{k=1}^K \left(\frac{(g_{i,j}^k - \mu_m^k)^2}{2(\sigma_m^k)^2} + \ln(\sqrt{2\pi}\sigma_m^k) \right) \right\}$$

The total energy E can be minimized introducing temperature and simulated annealing using a sampling method called Gibbs Sampler, because the posterior distribution is also Gibbs distributed. The most probable states under the Gibbs distribution are found by modeling the posterior distribution as another physical system. The temperature in the physical system is gradually reduced isolating the corresponding

low energy states ("annealing").

The probability of randomly changing a label is given by the probability of the feature data being of class m using the Gibbs distribution:

$$P(F = m) = e^{-\frac{E_{m=F_{i,j}}}{T}}$$

The global temperature T determines the sensitivity to energy variations and is reduced over time.

$$T(k) = \frac{c}{\log(1+k)}$$

is the temperature at simulated annealing time k , where c is a constant. Another cooling method is the fast cooling

$$T(k) = c \cdot e^{-gk},$$

where k is the simulated annealing time and g is a constant.

High temperatures allow many stochastic changes, decreasing as well as increasing the energy function. Lowering the temperature reduces the probability for a significant energy increase and forces the energy towards a local minimum. Beginning at high temperature a global optimum can be achieved, because significant changes can jump out of local minima in order to find a better optimum.

Minimizing the energy function requires estimates of μ and σ . Supervised implementations use training data to estimate μ and σ . The EM algorithm can be used like in the GMM to estimate μ and σ unsupervised. The EM algorithm given the MRF model using simulated annealing works as follows [16]:

1. Randomly initialize the segmentation image F .
2. E-Step: For all classes m estimate mean and standard deviation from the feature data $g_{i,j}$ based on the segmented image F :

$$\mu_m^k = \frac{1}{N_m} \sum_{i,j,F_{i,j}=m} g_{i,j}^k,$$

$$\sigma_m^k = \sqrt{\frac{1}{N_m-1} \sum_{i,j,F_{i,j}=m} (g_{i,j}^k - \mu_m^k)^2},$$

where N_m is the number of feature data being of class m

3. M-Step: Apply simulated annealing to minimize the energy function

(a) Set starting temperature and simulated annealing time $k = 0$.

(b) For each image data $g_{i,j}$

- For all classes calculate the total energy of being in class m

$$E_{i,j,m} = E_{R_{i,j,m}} + \alpha E_{F_{i,j,m}}$$

and the corresponding probability

$$P(F_{i,j} = m) = e^{-\frac{E_{i,j,m}}{T}}$$

- Generate a random number between 0 and 1. Accumulate the probabilities of all classes. Choose the label being the portion of the accumulated probability at the random probability.

(c) If the number of predefined iterations is reached, then stop, otherwise increment the simulated annealing time, decrease the temperature and repeat step b.

4. Repeat steps 2 and 3 until a stopping criterion is reached.

If a variable weighting factor [16] is used in the energy function, it has to be updated by the initialization of the simulated annealing in the M-Step.

3. Results

To test the image segmentation algorithms 30 infrared images are taken from the ReuSe [17] database. The ship area is segmented manually for calculating two simple quality measures. One is the classification rate, which is defined by

$$P_c = \frac{\text{no. classified pixels in the segment}}{\text{no. pixels in the segment}}.$$

The other measure is an error rate, which is given by

$$P_e = \frac{\text{no. misclassified pixels}}{\text{no. classified pixels}}.$$

From these measures the segmentation quality rate is derived by

$$P_s = P_c - P_e.$$

Except the stopping criteria, parameters are only necessary for the Fuzzy C-means and the MRF model. For the Fuzzy C-means clustering

the fuzzy factor $m = 2$ is chosen [10]. The MRF model has the following variable weighting factor [18]:

$$\alpha(t) = \beta \cdot (c_1 \cdot \gamma^t + c_2),$$

where β , γ , c_1 and c_2 are constants and t is the t -th EM iteration. Furthermore $\beta = 0.1$, $\gamma = 0.1$, $c_1 = 80$ and $c_2 = 1/K$ are taken, where K is the number of classes. Label energy E_R is determined once by a four neighborhood with $\beta = 1$ and by a eight neighborhood with $\beta = 0.5$. The simulated annealing is applied for 100 iterations, while the temperature decreases at each iteration from 15 to 0.1 using the fast cooling method.

Generally the algorithms stop if the estimated values converge. Only the MRF model works with a fixed number of EM iterations (25). Because of their statistical nature the energy could vary even if the clusters do not change significantly.

The effect of wavelet-based denoising of the infrared images before applying the clustering algorithms is investigated. Accordingly to the denoising results in [19] the BayesShrink [20] method is chosen. The denoising algorithm uses dyadic wavelet filters [21] [22] with $r = 1$ and $p = 2$ in 3 scales.

One problem in unsupervised image segmentation using clustering is the number of clusters in the image. Visually the three regions ship, sky and water can be distinguished, but sometimes there are more than three regions in the image. So some post-processing steps are necessary. After extracting the cluster with the greatest mean, its binary representation is labeled using sequential labeling [23]. The biggest cluster detected besides the background (which is usually the biggest) is assumed to be the ship. Heuristics are used to determine how likely the number of clusters fit to the image. If the number of clusters does not fit to the image, the number of clusters is incremented and the clustering process is repeated. The algorithm stops, if the number of clusters fits to the image or a maximum number of 11 clusters is reached.

The experimental results (Table. I) show three effects. Denoising improve the segmentation results of K-means and Fuzzy C-means clustering achieving about 5% higher segmentation rates. The segmentation quality of Gaussian Mixture Models and Markov Random Fields is decreased,

Segmentation Model	with denoise		without denoise	
	P_s	$\#\{P_s > 80\}$	P_s	$\#\{P_s > 80\}$
K-means Euclidean	66.33	10	74.00	12
K-means Mahalanobis	74.03	17	79.64	21
Fuzzy C-means Euclidean	65.67	12	70.36	11
Fuzzy C-means Mahalanobis	76.54	19	80.36	19
Gaussian Mixture Model	66.97	14	65.81	15
Markov Random Field 4 Neighborhood	78.32	20	77.67	19
Markov Random Field 8 Neighborhood	81.11	22	78.14	21

TABLE I

SEGMENTATION RATES AND NUMBER OF SEGEMENTED IMAGES WITH SEGMENTATION RATES HIGHER THAN 80 FOR DIFFERENT MODELS

because of smoothing effects by denoising. Comparing the distance models for K-means and Fuzzy C-means clustering the Mahalanobis distance delivers between 5% and 10% better results. Finally the Markov Random Field model with a 8 neighborhood achieves the highest segmentation rate of all tested segmentation procedures. Fig. 2 shows an example of MRF segmentation result. With the visual comparison of the segmentation results 80 is assumed as the lowest acceptable segmentation rate (Fig. 3). Thus there are only 22 acceptable segmented ships using the MRF model and 21 using K-means clustering. All segmented images in the MRF model with estimated regions greater than 4 were not accepted, while for instance using K-means clustering there are accepted segmentation results up to 11 regions. Therefore the model fitting procedure is of general nature and not optimized for the different clustering algorithms. Consequently optimized model fitting procedures for the different clustering algorithms can further improve the results.

The computing times are different for the clustering algorithms. K-means with Euclidean distance model is the fastest algorithm, while the Markov Random Field model with 8 neighborhood is the slowest of the tested algorithms. A segmentation with 3 regions takes an average of about 1000 seconds using the Markov Random Field model with 8 neighborhood on a 3Ghz Xeon Processor, but only 1 second with the K-means algorithm. The computing time increases linearly with the number of estimated regions, which is especially a problem for the MRF model. This emphasizes the necessity of optimal estimated regions again.

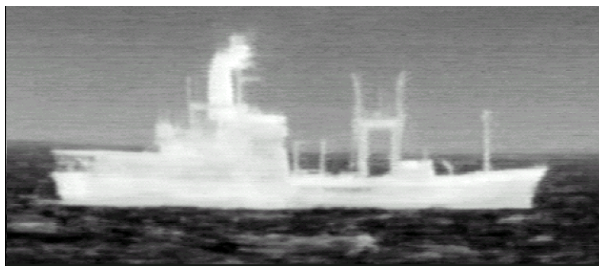
4. Conclusion

The target of our investigations was to find the best clustering algorithm to segment naval infrared images. It has been shown that the Markov Random Field model delivers the best results for our test images, but has also the highest computational demand. For integration in a ship contour extraction system the algorithm has to be accelerated. Therefore our future work will be focused on the acceleration of the MRF using parallel processing or special processors.

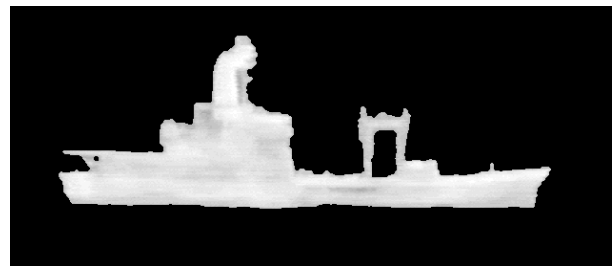
References

- [1] P. J. Withagen, K. Schutte, A. M. Vossepoel and M. G. Breuers: "Automatic classification of ships from infrared (FLIR) images", *Proceedings of SPIE Signal Processing, Sensor Fusion, and Target Recognition VIII*, Vol. 3720, No. 19, Jul. 1999.
- [2] J.A. Alves: "Recognition of Ship Types from an Infrared Image using moment invariants and neural networks", *Master Thesis*, Naval Postgraduate School, Monterey, Mar. 2001.
- [3] W. Meier and H.D. vom Stein: "Statistical Analysis of Infrared Image Sequences", *ISPRS Symposium: Spatial Information from Digital Photogrammetry and Computer Vision*, Munich, pp. 562-568, 1994.
- [4] W. Smeets: "Ein modellbasierter regionenorientierter Ansatz zur Segmentation von Infrarot-Bildsequenzen", *PHD Thesis*, UniBwH, Hamburg, 1995.
- [5] Y. Boykov and M.P. Jolly: "Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images", *Proceedings of 8th International Conference on Computer Vision*, Vol. I, pp. 105-112, 2001.
- [6] J. B. MacQueen: "Some Methods for classification and Analysis of Multivariate Observations", *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, University of California Press, Vol.1, pp. 281-297, 1967.
- [7] J.C. Dunn: "A Fuzzy Relative of the ISODATA Process and Its Use in Detection Compact Well-Separated Clusters", *Journal of Cybernetics*, Vol 3, Issue 3, pp. 32-57, 1974.
- [8] J.C. Bezdek: "Pattern Recognition with Fuzzy Objec-

- tive Function Algorithms", Plenum Press, New York, 1981.
- [9] L.A. Zadeh: "Fuzzy Sets", *Information and Control*, Vol 8, pp. 338-352, 1965.
- [10] Y.W. Lim and S.U. Lee: "On the Color Image Segmentation Algorithm based on the Thresholding and the Fuzzy C-means Techniques", *Pattern Recognition*, Vol. 23, No. 9, pp. 935-952, 1990.
- [11] D.E. Gustafson and W.C. Kessel: "Fuzzy Clustering with a Fuzzy Covariance Matrix", *Proceedings of the 17th IEEE Conference on Decision and Control*, pp. 761-766, San Diego, 1979.
- [12] A.P. Dempster, N.M. Laird and D.B. Rubin: "Maximum likelihood from incomplete data via the EM algorithm", *Journal Royal Statistics Society*, Vol. 39, No. 1, pp. 1-21, 1977.
- [13] J.A. Bilmes: "A Gentle Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models", *Technical Report*, TR-97-021, U.C. Berkeley, April 1998.
- [14] S. Geman and D. Geman: "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images", *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 6, No. 6, Nov. 1984.
- [15] Z. Kato, J. Zerubia and M. Berthod: "Unsupervised parallel image classification using Markovian models", *Pattern Recognition*, Vol. 32, No. 4, pp. 591-604, 1999.
- [16] H. Deng and D. A. Clausi: "Unsupervised Image Segmentation Using A Simple MRF Model with A New Implementation Scheme", *Proceedings of the 17th International Conference on Pattern Recognition*, Vol. 2, pp. 691-694, Nov. 2004.
- [17] W. Meier, K. Haese and H.D. vom Stein: "An integrated system for infrared sequence analysis and classification", *Signal and Image Processing - SIP 95*, Las Vegas, pp. 20-23, Nov. 1995.
- [18] J. Tran: "Segmentation of Dermatological Images using Mixture Models and Markov Random Fields", *Honours Thesis*, Monash University, Vitoria, Nov. 2005.
- [19] D. Wippig, B. Klauer and H. Ch. Zeidler: "Denoising of Infrared Images by Wavelet Thresholding", *Proceedings of the International Conference on Industrial Electronics, Technology and Automation - IETA 05*, Bridgeport, Dec. 2005.
- [20] S.G. Chang, B. Yu and M. Vetterli: "Adaptive Wavelet Thresholding for Image Denoising and Compression", *IEEE Transaction on Image Processing*, Vol. 9, No. 9, pp. 1532-1546, 2000.
- [21] S. Mallat and S. Zhong: "Characterization of Signals from Multiscale Edges", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14, No. 7, pp. 710-732, 1992.
- [22] I. Koren and A. Laine: "A Discrete Dyadic Wavelet Transform for Multidimensional Feature Analysis", *Time-Frequency and Wavelet Transform in Biomedical Engineering*, M. Akay (Editor), IEEE Press, New York, 1997.
- [23] W. Burger and M. J. Burge: "Digitale Bildverarbeitung - Eine Einfhruung mit Java und ImageJ", eXamen.press, Springer, Berlin, 2005.



(a)

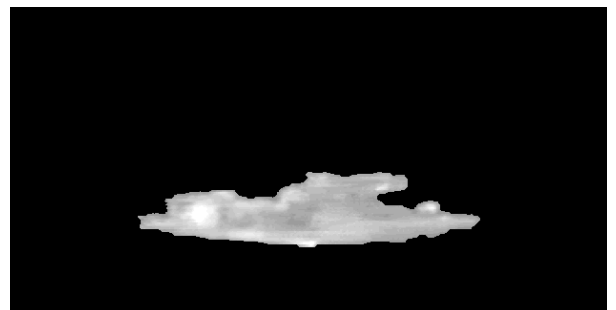


(b)

Fig. 2. (a) Infrared image, (b) MRF8 segmentation result ($P_s = 90.56$)



(a)



(b)

Fig. 3. (a) Infrared image, (b) MRF8 segmentation result ($P_s = 71.02$)