

A BORDER IRREGULARITY MEASURE USING HIDDEN MARKOV MODELS AS A MALIGNANT MELANOMA PREDICTOR

B. S. Aribisala* and Ela Claridge †
School of Computer Sciences
The University of Birmingham
Birmingham B15 2TT, U.K.

ABSTRACT

Malignant melanoma, a skin cancer, manifests itself as a dark lesion, most often with an irregular boundary. The degree of irregularity is an important diagnostic indicator. This paper presents a new measure of irregularity using Hidden Markov Models (HMMs) based on the Weibull probability distribution. The measure was tested on 98 skin lesions of which 16 were malignant melanoma. The ROC analysis showed that the measure is 82% sensitive and 82% specific in discriminating the malignant and benign lesions. These results compare favourably with other measures and indicate that HMM captures some distinguishing features in the boundary of malignant lesions.

KEY WORDS

Irregularity measure, Hidden Markov Models, Malignant melanoma, Weibull distribution, Skin lesions

1 Introduction

Melanoma is a malignant tumour of melanocytes. The tumour initially starts from the upper skin layer (epidermis) and later invades the dermis below. The survival rate for patients is inversely proportional to the depth of the tumour. Early detection of melanoma is the most important factor affecting the survival of a patient.

Malignant melanoma can be characterised using some physical features such as shape, edge, colour and surface texture. The border irregularity of pigmented skin lesions was identified by Keefe et al.[1] as the most significant diagnostic factor in clinical diagnosis of melanoma. Research by Morris Smith [2] revealed that “irregularity” is one of the major vocabulary terms used for describing border of the malignant melanoma in medical textbooks. The same research also showed that the clinicians place a significant emphasis on border irregularity when describing malignant melanoma. These findings coupled with the fact that border irregularity is one of the major features in the seven point checklist used for computing a “suspiciousness” score for skin lesions [3] indicate that border irregularity is a very important factor in the diagnosis of malignant melanoma.

It has been empirically discovered that clinicians have difficulties in visually assessing border irregularity of skin lesion outlines and that their assessments are not invariant to reflection and rotation [2, 4]. Much research on quantitative measures of irregularity has been carried out to overcome these shortcomings[2]. The most common approaches include the Compactness Index (e.g. [5]), Fractal Dimension (e.g. [6]) and measures based on radial distance (e.g. [7]). These methods are critically reviewed in a recent paper by Lee et al [8].

Claridge et al. [4, 9] proposed the use of fractal dimension (FD) and a modified fractal dimension called Structural Fractal Dimension (SFD). FD is not sensitive to the structural features such as indentations and protrusions [8]. SFD was designed to remove this drawback but its response to single indentations and protrusions, which can be indicative of early melanomas, was too weak to enable reliable classification. [8, 10].

Sigma Ratio and Indentation Irregularity Index, proposed by Lee and Atkins [8, 10], are non-linear measures derived by a curvature-scale filtering. They have been shown to be successful in detection of indentations and protrusions in the lesion border and have been the most successful algorithm to date for classifying skin lesions on the basis of their border irregularity.

The term “*Irregularity*” is intuitive and can express different meanings. If irregularity is to be quantified, it is necessary first to develop its formal definition, or at least provide its formal description. Five attributes of irregularity have been proposed [11]. One of these attributes which is of interest here is lack of predictability. The elements of a sequence corresponding to a regular shape or pattern are predictable, whereas in an irregular sequence they cannot be easily predicted. That is, the extent to which a sequence can be predicted may help us to determine how regular the sequence is.

In this paper we present a new measure of irregularity based on Hidden Markov Model. In contrast to the existing measures, the proposed measure is based on a formal criterion of irregularity outlined above. The model is trained on idealised regular lesion shapes represented as ellipses. If a given lesion is regular, its shape will be represented by the model. The degree to which a given lesion boundary conforms to the model is evaluated by computing the log-likelihood of the outline given the model. This measure

* Email: B.S.Aribisala@cs.bham.ac.uk

† Email: E.Claridge@cs.bham.ac.uk

estimates the extent to which a boundary sequence can be predicted by the model. This paper suggests that the log-likelihood value can be used to estimate border irregularity of skin lesions.

Section 2 presents a brief description of the Hidden Markov Models while section 3 presents the Markov Model for skin lesions. Section 4 describes the experiments on real lesion outlines. Results and discussion are in section 6. Finally, section 6 presents the conclusion.

2 Hidden Markov Model

A Markov model can be defined as a process that consists of a finite number of states N , and an $N \times N$ stochastic matrix with elements a_{ij} which gives the probabilities of transition from states i to j . A Hidden Markov Model (HMM) is a "composite" Markov model where each state has an associated probability density function $b_j(O_t)$ which gives the probability of the state j emitting a particular observation O_t at time t . The states are hidden and can only be observed through the emissions.

Let us denote a sequence of hidden states by $\{S_t\}$ and the associated sequence of observation by $\{O_t\}$ $t=1, \dots, T$ where T is the sequence length. HMM assumes that S_t depends only on S_{t-1} and that the observation O_t is independent of any other observations. This conditional independence permits us to define the joint probability distribution of the sequence of states and observations as

$$P(S_1, \dots, S_T, O_1, \dots, O_T) = P(S_1)P(O_1/S_1) \prod_{t=2}^T P(S_t/S_{t-1})P(O_t/S_t) \quad (1)$$

An HMM assumes that the state variables are discrete-valued, whereas the observation variables can be either discrete-valued or real-valued (continuous HMM). In this paper, we only consider a continuous HMM. In this instance HMM $P(O_t/S_t)$ can be modelled using different probability distributions such as Gaussian, Gamma, exponential or Weibull, or even mixture of probability distributions such as mixture of Gaussian (MOG). In this paper we model the observation sequence using Weibull distribution (see 3.2).

An HMM can be explicitly specified by using five parameters: N , M , A , B and π where N is the number of hidden states, M is the number of distinct observations, $A = a_{ij}$ is the transition probability distribution, $B = b_j(O_t)$ is the observation probability distribution matrix and $\Pi = \pi_i$ is the initial state distribution.

HMM has proved very useful in pattern recognition and pattern classification problems, for example in speech recognition [12, 13], handwriting recognition [14, 15], gesture classification [16], music classification [17] and EEG classification [18]. For further information on HMM see [12, 19, 20]

3 Skin Lesion Shape Model

3.1 Data Description and Problem Definition

Medical experts regard a skin lesion that is nearly circular or elliptical in overall shape as more likely to be normal than not [4]. In view of this we have taken the ellipse to be a shape model for a normal skin lesion and to represent the most *regular* instance of the lesion shape. The more irregular the lesion border, the less likely it is to conform to the model. Figure (1) shows examples of lesion outlines, (a) regular and (b) irregular.

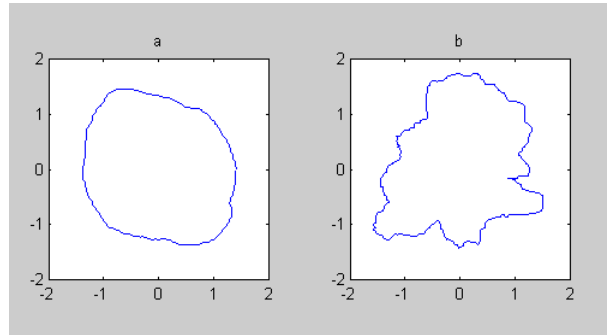


Figure 1. Examples of Lesion Outlines: (a) regular (b) irregular

A "distance" between the model and a given lesion data is a measure of the lesion's irregularity. If irregularity is framed as a lack of predictability (see section 1), it is easy to see that the points on the circumference of an ellipse can be easily predicted whereas points belonging to an irregular shape are not easily predictable.

The lesion border is represented as a sequence of (1D) radial coordinates in a polar coordinate system centred at the centre of gravity of the lesion. The coordinates constitute an observation sequence $O = O_1, O_2, \dots, O_M$ where O_i , $i=1, \dots, M$ is an i -th boundary point. A model λ describes a set of ellipses representing a normal skin lesion. Given the observation sequence O and the model λ we compute the probability $P(O/\lambda)$ which is the probability of the observation sequence O given the model λ . $P(O/\lambda)$ can equally be interpreted as the likelihood of generating the observation O when given the knowledge of model λ . Our hypothesis is that $\log(P(O|\lambda))$ (log-likelihood) should decrease with increase in border *irregularity*.

3.2 Choice of Probability Distribution for the Observation Sequence

In order to model the observation sequence we investigated the suitability of different probability distributions such as Gaussian, Gamma, Uniform, log-normal, exponential and Weibull (equation (2)). The investigation was carried out using a theoretical quantile-quantile (Q-Q) plot method

[21]. A Q-Q plot can be obtained by plotting the quantiles of a given sample data against the corresponding theoretical quantiles of a given probability distribution. The result of Q-Q plot is interpreted by examining the shape of the plot and the closeness of the plot to its best linear fit [21, 22]. Figures (2, 3 and 4) show the Q-Q plots for a typical skin lesion using Weibull distribution, Gaussian distribution and Gamma distribution respectively. Figure (2) shows that there are some outliers at the beginning of the plot, Figure (3) shows that there are outliers at the beginning and the end of the plot while Figure (4) shows that the plot skews to the left. These indicate that Gamma distribution does not represent the data well, and that the Gaussian distribution has more outliers than the Weibull distribution. These qualitative indicators were further confirmed by computing linear correlation coefficients for the three distributions and for all the lesion data. The mean and standard deviations were 0.977 (0.023) for Weibull distribution, 0.975 (0.025) for Gaussian distribution and 0.876 (0.066) for Gamma distribution.

As the Weibull distribution fits the data better than any of the two distributions, we have used it to model the observation sequence of the HMM for the skin lesions. The parameter α affects the shape of the distribution and parameter β modifies its scale.

$$f(y, \alpha, \beta) = \frac{\alpha}{\beta^\alpha} y^{\alpha-1} \exp\left(-\left(\frac{y}{\beta}\right)^\alpha\right), \alpha, \beta > 0 \quad (2)$$

4 Experiments

Experimental data consisted of 98 skin lesions of which of 16 were histologically confirmed cases of melanoma and the remaining 82 were benign lesions [23]. The radial coordinates corresponding to lesion boundary were extracted using a boundary modelling technique [24]. The data was normalised by subtracting the mean and dividing by standard deviation to make it scale invariant.

A Hidden Markov Model of the regular (elliptical) boundary was trained on a small set of ellipses, each represented by a sequence $y = y_1, y_2, \dots, y_M$ of normalised radial coordinates. The ratio of the major to minor radii of all the ellipses used is one. A model λ developed through this training is assumed to capture the essential characteristics of all similar y -s. Training employed an expectation maximisation method based on Baum-Welch algorithm [12, 19, 20]. The algorithm estimates the optimum values of the shape parameter α and scale parameter β of the Weibull distribution (equation (2)) by adjusting and re-estimating the parameters of HMM such that likelihood $P(M/y)$ is maximised. A fully connected HMM structure with 3 hidden states was used throughout.

Once the model has been trained, the testing procedure was carried out by computing $\log(P(O/\lambda))$ which gives the log-likelihood (probability) of generating the observation sequence O given the model λ . This was done for all the 98 cases considered. In order to further investigate

the appropriateness of the use of Weibull distribution in modelling the lesion boundary data, the experiments were repeated with the HMM trained using Gaussian probability distribution in place of Weibull distribution.

To assess the discriminatory power of the log-likelihood measure as a melanoma classifier, the ROC analysis [25] was performed for both HMMs (i.e Weibull and Gaussian distribution trained).

One interesting question, not answered through the above experiments, was whether any of the computed measures corresponds to the human perception of the border irregularity in the skin lesions. To this end an experimental survey was carried out. 20 skin lesion outlines randomly selected from the full data set were given to 23 subjects, none of whom had medical training. The subjects were asked to rank the outlines based on their degree of irregularity. The level of agreement among the subjects was evaluated using rank correlation r_s based on Kendall coefficient of concordance W [26]. The value of r_s ranges from 0 (no agreement) to 1 (perfect agreement).

To test a “default” hypothesis, that irregularity simply depends on the magnitude of variations along the boundary, the standard deviation was computed for all the lesion outlines.

The results were analysed as follows. The Spearman coefficient of correlation was determined for each pair of the three irregularity estimations: the log-likelihood measure, the visual assessment (using the average ranking from the 23 subjects), and the standard deviation for the selected 20 lesions. Finally, the relationship between all three estimations was examined using multiple linear regression analysis.

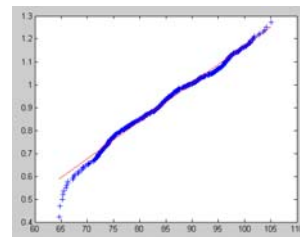


Figure 2. Quantile-Quantile plot for pigmented skin lesion data using Weibull distribution

5 Results and Discussion

The ROC analysis of the log-likelihood measure as a melanoma classifier showed 75% sensitivity and 76% specificity when the observation sequence was modelled using Gaussian probability distribution and 82% sensitivity with 82% specificity when using Weibull probability distribution. Standard deviation gave sensitivity of 77% and specificity of 78%. Figure (5) shows the ROC plots for

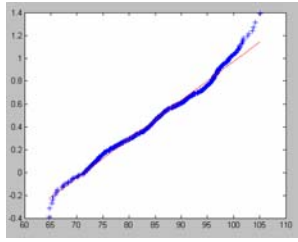


Figure 3. Quantile-Quantile plot for pigmented skin lesion data using Gaussian distribution

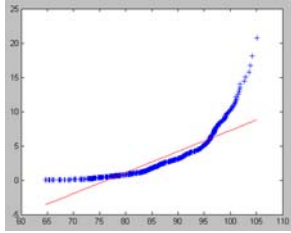


Figure 4. Quantile-Quantile plot for pigmented skin lesion data using Gamma distribution

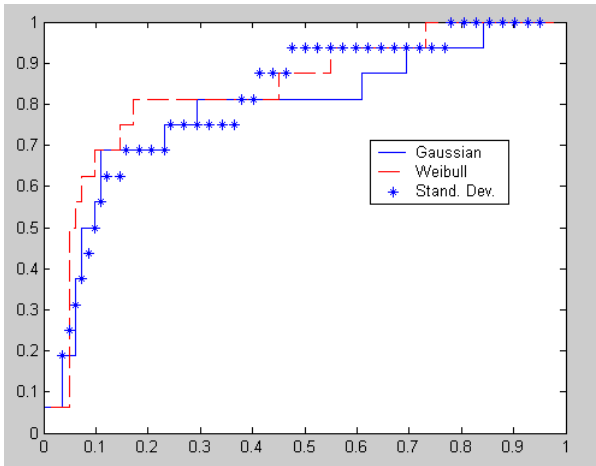


Figure 5. ROC curves for the log-likelihood measure using Gaussian distribution in the HMM (solid line), using Weibull distribution in the HMM (dashed line); and for the standard deviation (asterisks) - as melanoma predictor

all three cases. On the basis of these results we have concluded that boundary sequences corresponding to abnormal lesions are less predictable (i.e. more irregular) when assessed using the Weibull probability distribution model than when using the Gaussian model of a regular (assumed normal) boundary. The departure from the Gaussian model of the regular lesion as measured by log-likelihood is no more discriminatory than a simple standard deviation of the

boundary data.

In the experiments examining the perception of irregularity, the coefficient of concordance W was 0.886, indicating good agreement between all 23 subjects. The assessment of irregularity by the subjects correlated relatively poorly with both the Weibull based irregularity measure (Spearman correlation coefficient of 0.568) and standard deviation (0.466). Correlation was very poor between the Weibull based irregularity measure and standard deviation (0.229). The multiple linear regression analysis for all three tests showed equally poor correspondence (0.228). These results suggest that humans have similar way of assessing shape irregularity, but the human notion of irregularity may not be the same as measured by the indicators investigated in this paper. One possible interpretation of these results is that the features of a lesion boundary which makes it *look* irregular are not necessarily those associated with the irregularities present in abnormal lesions. A suitable computational irregularity measure may thus be a better abnormality predictor. This interpretation is only tentative because, first, the assessment of irregularity by clinicians is most likely to be different than those with no medical training; and second, they are drawn based on a small number of subjects (23). However, they are suggestive enough to encourage further work on this problem.

We have compared the result of the proposed measure with one of the best published melanoma predictors based on irregularity, the Indentation Irregularity Index (III) [23, 27]. The comparison used the area under the ROC curve, which is a global measure commonly used to assess the overall predictive power of classification schemes. The III computed for a superset of the set of lesions used in our experiments has the area under ROC curve of 0.73 [23], whereas for the HMM based measure the area is 0.81. This indicates that the HMM-derived measure has a greater discriminatory power than the III index.

6 Conclusion

In this paper we have proposed a new measure of border irregularity for pigmented skin lesions based on Hidden Markov Models. The measure has been devised to quantify one of the attributes of irregularity, namely a lack of predictability. We have demonstrated that regular lesion boundaries can be modelled using a Hidden Markov Model assuming Weibull distribution. Irregular boundaries do not conform to this model, which results in the decrease of the log-likelihood of the boundary being represented by the model. ROC analysis of the log-likelihood as a malignancy predictor gave 82% of both sensitivity and specificity. This result shows that the model captures some distinguishing features in the boundary of malignant lesions and thus can contribute to lesion classification.

References

- [1] M. Keefe, D Dick, and R. Wakeel. A study of the value of the seven point checklist in distinguishing benign pigmented lesions from melanoma. *Clin. and Exp. Derm.*, 15:167 – 171, 1990.
- [2] J. D. Morris-Smith. *Characterisation of the Appearance of Pigmented Skin Lesions*. PhD thesis, The University of Birmingham, U.K., 1996.
- [3] R. Mackie. Malignant melanoma, a guide to early diagnosis. Technical report, University Glasgow, 1989.
- [4] E. Claridge, P.N. Hall, M. Keefe, and J.P. Allen. Shape analysis for classification of malignant melanoma. *J. Biomed. Eng.*, 14(3):229–234, 1992.
- [5] W.V. Stoecker, R. H. Moss, F. Ercal, and S.E. Umbaugh. Nondermatoscopic digital imaging of pigmented lesions. *Skin Research and Technology*, 1:7–16, 1995.
- [6] P.N. Hall, E. Claridge, and J.D.M. Smith. Computer screening for early detection of melanoma - is there a future. *British Journal of Dermatology*, 132:325 – 338, 1995.
- [7] D. Gutkowicz-Krushin, M. Elbaum, P. Szwaykowski, and A. W. Kopf. Can early malignant melanoma be differentiated from atypical melanocytic nevus by in vivo techniques ? *Skin Research and Technology*, 3:15–22, 1997.
- [8] T. K. Lee, D. I. McLean, and M. S. Atkins. Irregularity index: A new border irregularity measure for cutaneous melanocytic lesions. *Medical Image Analysis*, 7(1):47–64, 2003.
- [9] E. Claridge, J. D. Morris-Smith, and Hall P. N. Evaluation of irregularity in pigmented skin lesions against a consensus of expert clinicians. In *Proceedings of Medical Image Understanding and Analysis, Leeds, U.K.*, pages 85–88, 1998.
- [10] T. K. Lee and M.S. Atkins. A new shape measure for melanocytic lesion. In *Proceeding of Medical Image Understanding and Analysis, London, U.K.*, pages 25 – 28, 2000.
- [11] B.S. Aribisala. Computing irregularity for features in medical images. Technical report, The University of Birmingham, U.K., May 2003.
- [12] L. R. Rabiner. A tutorial on Hidden Markov Models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):275 – 286, 1989.
- [13] J. G. Wilpin and L.W. Rabiner. Automatic recognition of keywords in unconstrained speech using Hidden Markov Models. *IEEE Trans. on ASSP*, 38(11):1870–1878, 1990.
- [14] T. F. Sezgin. Genetic and HMM based approaches to freehand sketch recognition. sow.csail.mit.edu/2003/proceedings/SezginHmm.pdf, Last Viewed on 06/10/04.
- [15] M. Pechwitz and V. Maergner. HMM based approach for handwritten arabic word recognition using the IFN/ENITs. In *Database Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR)*, 2003.
- [16] N. Lin and C. L. Brian. Gesture classification using Hidden Markov Models and viterbi path counting. In *Proceeding: VIIth Digital Image Computing Techniques and Applications, Sydney*, 2003.
- [17] W. Chai and B. Vercoe. Folk music classification using hidden markov models. In *Proceedings of International Conference on Artificial Intelligence*, 2001.
- [18] H. Lee and S. Choi. PCA+HMM+SVM for EEG pattern classification. <http://www.postech.ac.kr/seungjin/publications/isspa03.pdf>, Last Viewed on 06/10/04.
- [19] J. Bilmes. What HMMs can do. Technical Report UWEET-2002-0003, Department of Electrical Engineering, University of Washington, 2002.
- [20] Y. Wu, A. Ganapathiraju, and J. Picone. Report for Baum-Welch re-estimation of Hidden Markov Model. Technical report, Institute for Signal and Information Processing, Department of Electrical and computer Engineering, Mississippi state University, 1999.
- [21] J. M. Chambers, W.S. Cleveland, B. Kleiner, and P.A. Tukey. *Graphical Methods for Data Analysis*. Wadsworth, 1983.
- [22] W. S. Cleveland. *Elements of Graphing Data*. Wadsworth, 1985.
- [23] T. K. Lee and E. Claridge. Predictive power of irregular border shapes for malignant melanoma. *Skin Research and Technology* (Accepted for Publication).
- [24] E. Claridge and A. Orun. Modelling of edge profiles in pigmented skin lesions. *Medical Image Understanding and Analysis 2002*, pages 53–5, 2002.
- [25] A.R. Webb. *Statistical Pattern Recognition*. John Wiley and Sons Ltd., 2003.
- [26] W. Hays. *Statistics, Chap. 19*. Holt, Rinehart and Winston, 1988.
- [27] T. K. Lee and M.S. Atkins. A new approach to measure border irregularity for melanocytic lesions. In *SPIE Medical Imaging, San Diego*, pages 668–675, 2000.