

Texture Features in the Classification of Melanocytic Lesions

Jukka Kontinen*, Juha Rönning* and Rona M. MacKie**

*Machine Vision and Media Processing Group,
Infotech Oulu and Department of Electrical Engineering,
University of Oulu
FIN-90570 Oulu, Finland, email: juko@ee.oulu.fi

**Department of Dermatology, Robertson Building,
University of Glasgow,
Glasgow G12 8QQ, England

Abstract

The use of different texture features for the classification of melanocytic lesions is studied in an attempt to develop a computerized method for the early detection of melanoma. The computer laboratory at the University of Oulu has a strong tradition in applications of computer vision to visual inspection for industrial quality control, and some of the methods learnt in these applications are being transferred and experimented with in this medical context. This includes the utilization of texture distributions for classification purposes.

To avoid the effect of different photographing systems, all the images are first converted to intensity images and then the lesion parts are divided into 32x32 rectangles in order to obtain the maximal number of non-overlapping samples. The divided images are then normalized by z-score transformation and texture feature distributions are counted for the rectangular samples and classified into melanoma and benign nevus with a k-nearest neighbour classifier. 78-99% of the test samples were found to be classified correctly, depending on the texture feature used.

1. Introduction

Skin cancers are the only cancer types that are visible from the early stages. Since the accuracy of clinical diagnosis of melanoma is only around 75% for very experienced dermatologists [1, 2] and even lower for non-dermatologists [3], computer vision methods are a tempting alternative as an aid in its early detection and as a means of acquiring and storing images of pigmented lesions and implementing teledermatological methods [4].

There has been a rapid increase in skin cancer, particularly malignant melanoma, the annual incidence rate in Finland having quadrupled within 25 years (from years 1958-60 to 1985-87) [5]. In Denmark it is estimated that there will be 1500 new cases of melanoma in the year 2000 compared with approx. 600 in the year 1990 [6]. The incidence of malignant melanoma is increasing faster than that of any other cancer [7]. Since there is no effective treatment for patients with melanoma metastases, the only chance of a cure is early detection [8].

Several investigations have been carried out recently into computer image analysis for the diagnosis of melanoma. Schindewolf *et al.* used scanned and digitized colour

slides of melanocytic lesions to create images for computer classification by colour and texture, having an accuracy of 92% [9]. Green *et al.* used a hand-held system with a video camera and colour frame grabber mounted in a microcomputer to obtain images for software analysis, extracting features relevant to the size, colour, shape and boundary of each lesion. The outcome was that 89% of pigmented lesions were correctly classified [10]. Ganster *et al.* introduced computerized analysis of images obtained from epiluminescence microscopy (ELM), and found 86% of the test images to be classified correctly [11].

Commonly used features in experiments arrived at the computerized diagnosis of melanoma are shape, colour, symmetry, area and structure. Most of these need accurate boundary detection which can be one of the most demanding and time-consuming parts of the work. We consider here the correlation of the texture features with melanoma, a method which does not require as accurate boundary detection as others. Boundary detection was not the main object of study at this point, however, and in practice the texture rectangles were assessed manually inside the lesion area.

2. Materials and Methods

Our data set consisted of 86 colour slides, which had been provided by the Department of Dermatology and Venereology at Oulu University Hospital and the Department of Dermatology at the University of Glasgow. All the test images were assigned histologically beforehand to two specified diagnostic groups, "melanoma" and "benign melanocytic nevus" (Fig. 1). All the images included a piece of measuring tape, by which they were scaled to the same size after digitization.

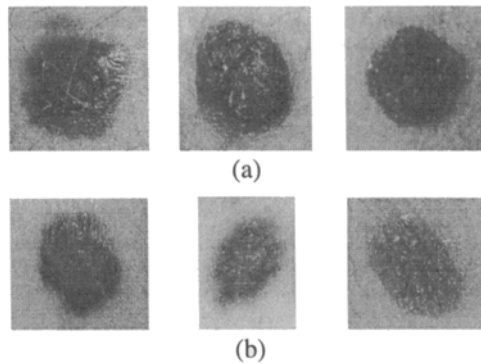


Fig. 1. Examples of (a) Melanoma images; (b) Benign nevus images.

2.1 Image Processing System

The image digitization was carried out using an Agfa Arcus scanner with slide deck, using 600 dpi resolution with 24 bit depth coding (8 bit/primary colour and 3 primaries Red, Green, Blue). All the image processing was done with Sun Sparcstation 20 workstations using image processing software developed in the Khoros software development environment [12].

2.2 Image Normalization

All the images were normalized with z-score transformation [13] in order to avoid the effect of different photographic systems, but before this they were converted to intensity images and the lesion parts were divided into rectangles in order to obtain maximal number of non-overlapping samples 32x32 pixels in size. The total of 86 images yielded 27 melanomas and 59 nevi, represented by 462 melanoma samples and 787 benign nevus samples and producing a data set of 1249 samples.

The images were normalized to have an average value of 120 and a standard deviation at 30. Since all the original intensity images had a deviation of less than 30, no loss of information occurred in the pre-processing.

2.3 Texture Features

The following texture features were used [14]:

Local Binary Pattern (LBP)

LBP is a two-level version of the texture unit method of Wang and He [15]. It provides a robust means of describing pure local binary patterns in a texture. The original 3 x 3 neighbourhood (Fig. 2a) is thresholded by the value of the center pixel. The values of the pixels in the thresholded neighbourhood (Fig. 2b) are multiplied by the weights given to the corresponding pixels (Fig. 2c). The result for this example is shown in Fig. 2d. Finally, the values of the eight pixels are summed to obtain the number of this texture unit (169). LBP is invariant in monotonic grey scale transformation.

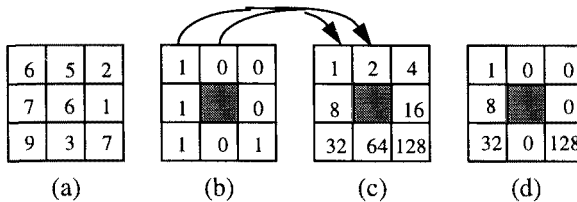


Fig. 2. Two-level version (LBP) of the texture unit.

The LBProt8 method is a rotation invariant version of LBP (Fig. 3). The binary values of the thresholded neighbourhood (Fig. 3a) are mapped into an 8-bit word in clockwise or counter-clockwise order (Fig. 3b), and an arbitrary number of binary shifts is made (Fig. 3c) until the word matches one of the 36 different patterns (Fig. 3d) of '0' and '1' that an 8-bit word can form under rotation.

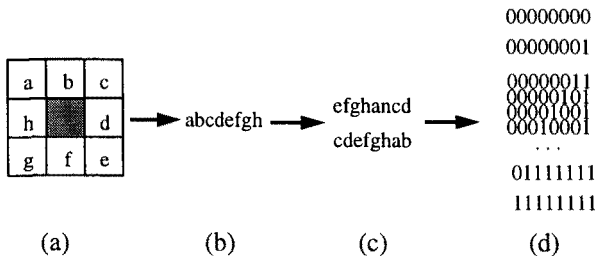


Fig. 3. Rotation invariant version of LBP.

Center-symmetric covariance measures

Center-symmetric covariance measures are abstract measures of texture pattern, grey-scale and the amount of local texture. A mathematical description of these measures computed for center-symmetric pairs of pixels in a 3x3 neighbourhood (Fig. 4) is presented in equations (1)-(8), where x_i refers to the grey level of pixel i and μ denotes the local mean.

x_2	x_3	x_4
x_1		x_1'
x_4'	x_3'	x_2'

Fig. 4. A 3x3 neighbourhood with 4 center-symmetric pairs of pixels.

$$SCOV = \frac{1}{4} \sum_i^4 (x_i - \mu)(x_i' - \mu) \quad (1)$$

$$SRAC = 1 - \frac{12 \left\{ \sum_i^4 (r_i - r_i')^2 + T_x \right\}}{m^3 - m} \quad (2)$$

where $m = n^2 - 1$

r_i refers to the rank of the grey level of pixel i in the ranked 3x3 neighbourhood of Fig. 4.

and

$$T_x = \frac{1}{12} \sum_j^l (t_j^3 - t_j) \quad (3)$$

where t_j = number of ties at rank r_i

l = number of ranks

$$VAR = \frac{1}{8} \sum_i^4 (x_i^2 + x_i'^2) - \mu^2 \quad (4)$$

$$BVAR = \frac{1}{16} \sum_i^4 (x_i + x_i')^2 - \mu^2 \quad (5)$$

$$SAC = \frac{SCOV}{VAR} \quad (6)$$

$$WVAR = \frac{1}{16} \sum_i^4 (x_i - x_i')^2 \quad (7)$$

$$SVR = \frac{WVAR}{BVAR} \quad (8)$$

Laws' texture energy measures

Laws' properties are derived from three simple vectors of length 3, $L3 \equiv (1,2,1)$, $E3 \equiv (-1,0,1)$, and $S3 \equiv (-1,2,-1)$ [16]. By multiplying the column vectors by row vectors of the same length, we obtain Laws' 3x3 masks. The four zero-sum 3x3 masks used here are shown in Fig. 5, where L3S3 and L3E3 are spot and edge detectors, respectively.

-1	0	1	-1	-2	-1	-1	2	-1	-1	-2	-1
-2	0	2	0	0	0	-2	4	-2	2	4	2
-1	0	1	1	2	1	-1	2	-1	-1	-2	-1
L3E3			E3L3			L3S3			S3L3		

Fig. 5. The four Laws' 3x3 masks.

Grey level differences

Four other measures based on grey level differences are also used. DIFFX and DIFFY are histograms of absolute grey level differences between neighbouring pixels computed in horizontal and vertical directions. DIFF2 accumulates absolute differences in horizontal and vertical directions and DIFF4 in all four principal directions in a single histogram, providing rotation invariant texture measures.

2.4 Distribution Classification

The feature distributions of the 32x32 samples are estimated by means of a histogram with a fixed number of bins, In this case 32.

The feature space was quantified by summing the feature distributions for ten benign nevus and ten melanoma images in a total distribution which was divided into 32 bins, each having an equal number of entries. In this way the highest resolution of the quantification is used where the number of entries is largest, and vice versa. Since LBProt8 has a discrete set of output values, no quantification is needed, but instead the operator outputs are accumulated directly into the histogram.

The classification was carried out with a kNN classifier using a k value of 3 and the leave-one-out method, i.e. when one sample was classified the other 1248 samples served as models. The distance between the sample and each model was computed and the sample was assigned to the class to which the majority of the 3 nearest models belonged. The distance metric used for measuring the similarity between distributions was histogram intersection [17]:

$$HIS = 1 - \sum_{i=1}^n \min(S_i, M_i) \quad (9)$$

where S_i and M_i are the respective sample and model probabilities at bin i and n is the number of bins.

3. Results

The classified lesion samples were assigned to two classes, "melanoma" and "benign nevus". Table 1 shows the total error for all individual texture features. Note that the error shows the percentage of the 32x32 samples misclassified.

Table 1: Classification errors for all individual texture features

Feature	Total error [%]	Feature	Total error [%]	Feature	Total error [%]
LBProt8	9.05	BVAR	12.33	DIFFX	1.28
SCOV	13.61	WVAR	19.46	DIFFY	1.52
SAC	11.93	L3E3	18.65	DIFF2	1.36
SRAC	10.09	E3L3	16.81	DIFF4	1.52
SVR	11.93	L3S3	21.54		
VAR	19.86	S3L3	17.21		

The fact that the number of test images were not evenly divided between the two photographing systems may have a minor affect on the results, but they still show a strong correlation between the DIFF features and melanoma. The total error of 1.28% for the DIFFX feature means only 16 samples misclassified out of 1249. The DIFF features seems to correlate with the variation in pigmentation within a lesion.

The center-symmetric covariance measures and the LBProt8 texture feature correlate with the texture pattern and the amount of local texture, and their correlation with the melanoma is more complicated than with DIFF features, although it still gives 80-91% accuracy in classification.

The Laws' texture energy measures are supposed to correlate with the spots and edges within a lesion. The number of different structural components is one of the four criteria in the ABCD rule of dermatoscopy introduced by Stolz *et al.* [18], which is used to discriminate between malignant melanoma and benign melanocytic nevus. The Laws' texture features seem to give steady classification results, although the poorest of all the texture features used here, with 78-83% accuracy.

We also examined the discriminative power of feature combinations. Several fea-

tures were combined by summing the similarity scores computed for each individual texture measure (Table 2).

The improvement in the classification results when combining different types of texture features is clear, and we can assume that these combinations measure the lesion in a more robust way than any of the features alone.

Table 2: Classification errors for different combinations of features

Features	Total error [%]
LBProt8, SVR, S3L3	6.00
LBProt8, VAR, E3L3	5.76
LBProt8, BVAR	4.24
LBProt8, DIFFX	0.88
LBProt8, DIFF2	0.96

4. Discussion and Conclusion

The above results clearly demonstrate the value of different texture features for the detection of melanoma, and it is easy to see from the results that texture features that measure grey level differences correlate especially well with melanoma lesions. This was to be expected, because one of the basic visual features that indicates melanoma is the variation in pigmentation within a lesion.

Our test images represented two distinct groups, benign nevus and melanoma, and no dysplastic nevus cases were included. We believe that in most cases, especially when the melanocytic lesion is at the transition between benign and malignant, the best discrimination is to be obtained by combining several texture features. A larger test image set, including a considerable number of dysplastic nevi, would enable the best combination of texture features to be found.

5. Acknowledgements

This research was supported by the Cancer Society of Finland and the Tauno Tönning Foundation. The authors gratefully acknowledge the expert assistance of Aarne Oikarinen and Leena Savolainen of the Department of Dermatology and Venereology at Oulu University Hospital and Timo Ojala of the Department of Electrical Engineering at the University of Oulu.

6. References

- [1] Grin, C.M., Kopf, A. W, Welkovich, B., et al. Accuracy in the clinical diagnosis of malignant melanoma, Archives of Dermatology, Vol. 126, June 1990, pp. 763-766.

- [2] Schmoeckel, C., Wagner-Grösser, G. und Braun-Falco, O., Klinische diagnostik initialer maligner melanome, *Der Hautarzt*, Vol. 36, 1985, pp. 558-562.
- [3] Rampen, H. J. and Rümke, P., Referral pattern and accuracy of clinical diagnosis of cutaneous melanoma, *Acta Dermato-Venereologica*, Vol. 68, 1988, pp. 61-64.
- [4] Stoecker, W.V., Moss, R.H., Ercal, F. and Umbaugh, S.E., Nondermatoscopic digital imaging of pigmented lesions, *Skin Research and Technology*, Nr 1, 1995, pp. 7-16.
- [5] Teppo, L., Ihosyöpa Suomessa, *Suomen Lääkärilehti*, 28/91 vsk 46, pp. 2598-2601.
- [6] Osterlind, A., Jensen, O. M., Hudpigmentering, naevi og sollys som aetiologiske faktorer til kutant malignant melanom, *Nordisk Medicin*, Vol. 195, 8-9/1990, pp. 223-229.
- [7] Ries, L.A.G, Hankey B.F., Miller B.A. et al. Cancer statistics review 1973-1988. Bethesda, Md: National Cancer Institute. Surveillance Epidemiology and End Results (SEER) data base. NIH publication Nr. 91-2789, 1991.
- [8] Friedman, R. J., Rigel, D. J., and Kopf, A. W., Early Detection of Malignant Melanoma: The Role of Physician Examination and Self-Examination of the Skin; CA-A Cancer Journal for Clinicians, Vol. 35, Nr 3, May/June 1985, pp. 130-151.
- [9] Schindewolf, T., Stolz, W., Albert, R., et al. Classification of melanocytic lesions with color and texture analysis using digital image processing, *Analytical and Quantitative Cytology and Histology*, Vol. 15, Nr 1, February 1993, pp. 1-11.
- [10] Green, A., Martin, N., Pfitzner, J., et al. Computer image analysis in the diagnosis of melanoma, *Journal of the American Academy of Dermatology*, Vol. 31, Nr 6, December 1994, pp. 958-964.
- [11] Ganster, H., Gelautz, M., Pinz, A., et al. (1995) Initial results of automated melanoma recognition, *Proceedings of the 9th Scandinavian Conference on Image Analysis*, Uppsala, Sweden, pp. 209-218.
- [12] The Khoros Group (1991) *Khoros User's Manual & Programmer's Manual*. University of New Mexico, USA.
- [13] Milligan, G.W. and Cooper, M.C., A study of standardization of variables in cluster analysis, *Journal of Classification*, Nr 5, 1988, pp. 181-204.
- [14] Ojala, T., Pietikäinen, M. and Harwood, D., A comparative study of texture measures with classification based on feature distributions, *Pattern Recognition*, Vol. 29, Nr 1, 1996, pp. 51-59.
- [15] Wang, L. and He, D.C., Texture classification using texture spectrum, *Pattern Recognition*, Vol. 23, 1990, pp. 905-910.
- [16] Laws, K.I., *Textured image segmentation*, Image Processing Institute, University of Southern California, Report 940, 1980.
- [17] Swain, M.J. and Ballard, D.H., Indexing Via Color Histograms, *Proceedings of Third International Conference on Computer Vision*, Osaka, Japan, December 4-7, 1990, pp. 390-393.
- [18] Stolz, W., Hölzel, D., Riemann, A., et al. (1991) Multivariate analysis of criteria given by dermatoscopy for the recognition of melanocytic lesions. In: *Book of Abstracts, Fiftieth Meeting of the American Academy of Dermatology*, Dallas, Tex: Dec. 7-12.