Quantitative assessment of tumour extraction from dermoscopy images and evaluation of computer-based extraction methods for an automatic melanoma diagnostic system

Hitoshi Iyatomi, Hiroshi Oka, Masataka Saito, Ayako Miyake, Masayuki Kimoto, Jun Yamagami, Seiichiro Kobayashi, Akiko Tanikawa, Masafumi Hagiwara, Koichi Ogawa, Giuseppe Argenziano, H. Peter Soyer and Masaru Tanaka

The aims of this study were to provide a quantitative assessment of the tumour area extracted by dermatologists and to evaluate computer-based methods from dermoscopy images for refining a computer-based melanoma diagnostic system. Dermoscopic images of 188 Clark naevi, 56 Reed naevi and 75 melanomas were examined. Five dermatologists manually drew the border of each lesion with a tablet computer. The inter-observer variability was evaluated and the standard tumour area (STA) for each dermoscopy image was defined. Manual extractions by 10 non-medical individuals and by two computer-based methods were evaluated with STA-based assessment criteria: precision and recall. Our new computer-based method introduced the region-growing approach in order to yield results close to those obtained by dermatologists. The effectiveness of our extraction method with regard to diagnostic accuracy was evaluated. Two linear classifiers were built using the results of conventional and new computer-based tumour area extraction methods. The final diagnostic accuracy was evaluated by drawing the receiver operating curve (ROC) of each classifier, and the area under each ROC was evaluated. The standard deviations of the tumour area extracted by five dermatologists and 10 non-medical individuals were 8.9% and 10.7%, respectively. After assessment of the extraction results by dermatologists, the STA was defined as the area that was selected by more than two dermatologists. Dermatologists selected the melanoma area with statistically smaller divergence than that of Clark naevus or Reed naevus ($P=0.05$). By contrast, non-medical individuals did not show this difference. Our new computer-based extraction algorithm showed superior performance (precision, 94.1%; recall, 95.3%) to the conventional thresholding method (precision, 99.5%; recall, 87.6%). These results indicate that our new algorithm extracted a tumour area close to that obtained by dermatologists and, in particular, the border part of the tumour was adequately extracted. With this refinement, the area under the ROC increased from 0.795 to 0.875 and the diagnostic accuracy showed an increase of approximately 20% in specificity when the sensitivity was 80%. It can be concluded that our computer-based tumour extraction algorithm extracted almost the same area as that obtained by dermatologists and provided improved computer-based diagnostic accuracy. *Melanoma Res* 16:183–190 © 2006 Lippincott Williams & Wilkins.

Melanoma Research 2006, **16**:183–190

Keywords: automatic tumour extraction, Clark naevus, dermoscopy, melanoma, Reed naevus, remote diagnosis

Department of Dermatology, School of Medicine, Keio University, Tokyo, Japan

Correspondence and requests for reprints to Hitoshi Iyatomi Dr Eng, Department of Dermatology, School of Medicine, Keio University, 35 Shinanomachi, Shinjuku-ku, Toyko, Japan
Tel: +81-3-5363-3823; fax: +81-3-3351-6880; e-mail: iyatomi@k.hosei.ac.jp

Received 1 May 2005 **Accepted** (after revision) 17 January 2006

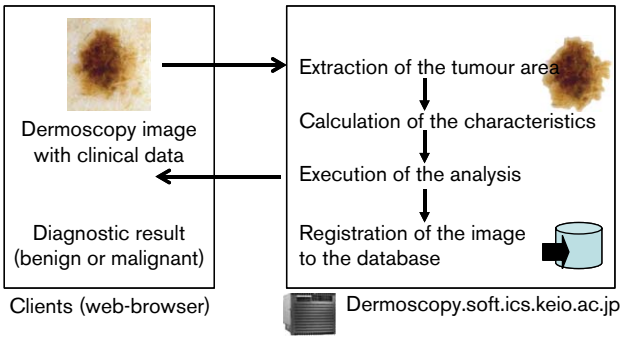
Introduction

The incidence of malignant melanoma has increased in most parts of the world over the past few decades. For example, in the northern part of Australia, the incidence is now approaching 50 cases per 100 000 population [1]. Although advanced melanomas are often fatal, early-stage lesions are mostly curable if resected, especially when free of metastasis. In particular, patients with melanomas equal to or less than 0.75 mm in thickness have a good prognosis, and the 5-year survival rate is reported to be more than 93% [2–5]. Therefore, the early detection and correct diagnosis of melanoma at an early stage are important for a reduction in the melanoma-related mortality rate.

Discrimination between early melanomas and Clark naevi is often difficult with the naked eye, even for expert dermatologists, especially when the lesions are still small. Dermoscopy, or epiluminescence light microscopy, has been developed to establish a correct diagnosis of pigmented skin lesions (PSLs) [6]. However, dermoscopy is often subjective and associated with low reproducibility

0960-8931

Fig. 1



Protocol used by the web-based diagnostic system. When the server receives a dermoscopy image and clinical data, it performs tumour area extraction, calculation of the tumour characteristics and judgement based on an artificial neural network classifier. The server then registers the data to the database and finally sends back the diagnostic results to the client.

and potential errors in the diagnosis of PSLs [1]. Computer analyses of PSLs could overcome the subjectivity of dermoscopy. Several groups have already developed automated analysis software in order to overcome these problems and have reported high sensitivity and specificity [7–12].

On the basis of this background, we established an automated analysis program and developed a web-based screening system (http://dermoscopy.soft.ics.keio.ac.jp/index_eng.html) [13]. Currently, a new and more stable version of the system is operational. Figure 1 illustrates the protocol involved in our system. When the server receives a dermoscopy image together with the related clinical data, it first performs tumour area extraction, calculation of the lesion characteristics and judgement based on an artificial neural network classifier. The server then registers the data into the database and finally sends back the diagnostic results (benign or malignant and a plausibility value that is calculated by the neural network classifier) to the client user.

Thus, to construct an automated melanoma diagnostic system, extraction of the tumour area from a given dermoscopy image and quantification of the lesion features are essential steps. Because the results of these steps directly affect the accuracy of the diagnosis, the detection of the appropriate tumour area is one of the most important steps.

Several techniques for tumour area extraction from dermoscopy images have been reported since the late 1990s. These techniques can be categorized into thresholding operation [14,15], segmentation operation [16,17] and combinations of the two [18,19]. Green *et al.* [14] achieved 83.8% correct extraction from 204 dermoscopy images. Later, Ganster *et al.* [18] combined the thresholding and segmentation techniques and achieved around 96% correct extraction from 4000 epiluminescence microscope images. However, these results were evaluated without detailed verification, and only a few quantitative evaluations have so far been performed. In our preliminary study, we confirmed that the results of conventional computer-based extraction were generally less than the tumour area recognized by dermatologists [20]. On the other hand, Schmid-Saugeon *et al.* [19,21] statistically evaluated the manual extraction of 25 images by five dermatologists, and demonstrated that even expert dermatologists could not reproduce the same results. They concluded that the validation of boundary detection could not be performed with a result from a single dermatologist. Although we agree with their arguments, there is still a need to define a suitable reference as a standard to develop an appropriate computer-based tumour area extraction system. Although extraction by dermatologists is not reproducible and cannot be regarded as the gold standard, the average or a comparable result can be considered as a reliable reference when a large number of results are provided. For these reasons, it is necessary to conduct a quantitative evaluation of manual extraction by dermatologists and to define the standard tumour area (STA). In the present study, we evaluated the results of extraction by non-medical individuals and by two computer-based methods based on the STA.

Materials

Dermoscopic digital images of PSLs were collected from two university hospitals (University of Naples, Italy, and University of Graz, Austria) and stored in JPEG format. All cases were diagnosed on the basis of histopathological examination of biopsy material. The digital images were selected if they fulfilled the following three conditions: (1) the images contained few hairs; (2) acral and mucosal lesions were excluded; (3) all parts of the pigmented lesion were included inside the image frame. A total of 188 cases of Clark naevi, 56 cases of Reed naevi and 75 cases of melanoma, including 23 cases of melanoma *in situ*, fulfilled the above criteria (total of 319 cases) and were included in the study.

Methods

Definition of the STA and evaluation criteria

As it is necessary to define the canonical tumour area, we established the gold standard of the tumour area, that is the 'standard tumour area' (STA), after assessment of manual extraction by five expert dermatologists.

Firstly, each of the five dermatologists (average experience of 11 years) manually specified the area of each tumour from 319 dermoscopy images by tracing the outline of the tumour images on a tablet computer.

Secondly, we used two candidates for the STA as follows: 'average area', the average of the extraction results by the five dermatologists; 'major area', the area selected by more than two dermatologists. Thirdly, the standard deviation (SD) of these areas and the area extracted by 10 non-medical individuals were examined. Finally, we assessed the above results and defined the STA from the two candidates (average or major area).

After STA had been defined, we introduced the evaluation criteria of the extraction performance for non-medically trained individuals and computer-based methods, namely precision and recall [22]. The precision and recall are defined as follows (Fig. 2): precision = (correctly extracted area)/(extracted area); recall = (correctly extracted area)/STA. Note that the 'correctly extracted area' represents the intersecting parts of the STA and the extracted area.

Evaluation of the area extracted by non-medical individuals

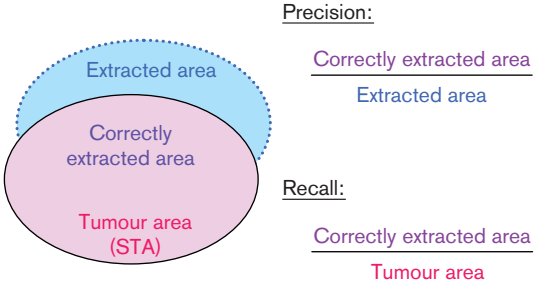
Hand-operated extraction was examined by 10 non-medically trained individuals. In 3190 images (319 images × 10), each extraction was estimated by precision and recall. The SD of the area extracted by 10 non-medical individuals was also compared with that of dermatologists.

Evaluation of the area extracted by computer-based methods

Two computer-based tumour area extraction methods were examined by precision and recall. The first computer-based method was a conventional thresholding-based method [23] and the other was the following coordinated method [20].

The conventional method used a statistically optimum colour threshold to separate a given dermoscopy image into two areas. The threshold was decided in such a way that maximized the inter-group variance of the blue information of each pixel of the image, and the darker area (whose blue value was lower than the threshold) was extracted as the tumour area.

Fig. 2



Tumour area and extracted area. Precision and recall are the performance indices of tumour area extraction. Precision reflects the accuracy of the extracted area and recall indicates how well the tumour area is extracted. STA, standard tumour area.

After assessment of the comparative experiments of the computer-based algorithm versus several dermatologists, the area extracted by the dermatologists was found to be generally larger than that obtained by several computer-based methods [20]. Hence, in our coordinated method, the temporarily extracted area was adjusted by a region-growing approach to bring it close to the area selected by dermatologists.

The coordinated algorithm was applied as follows: (1) the thresholding tumour area extraction method [23] was performed; (2) the isolated area was separated and small regions were integrated; (3) the region that fulfilled certain conditions (Fig. 3) was regarded as the tumour area and was extracted; (4) the extracted area was adjusted using the region-growing approach so as to bring the area close to the STA.

Let us assume a pixel on the border of a tumour area (black pixel) and an $S \times S$ pixel window with the target pixel in its centre (Fig. 4). For every border pixel (brick pixel), the average intensities of the inside and outside of the tumour area, namely V_{in} and V_{out}, were calculated. In the following equation, we regarded the border as ambiguous and treated the $S \times S$ area as the tumour area. This process was continued until the location of the border became constant.

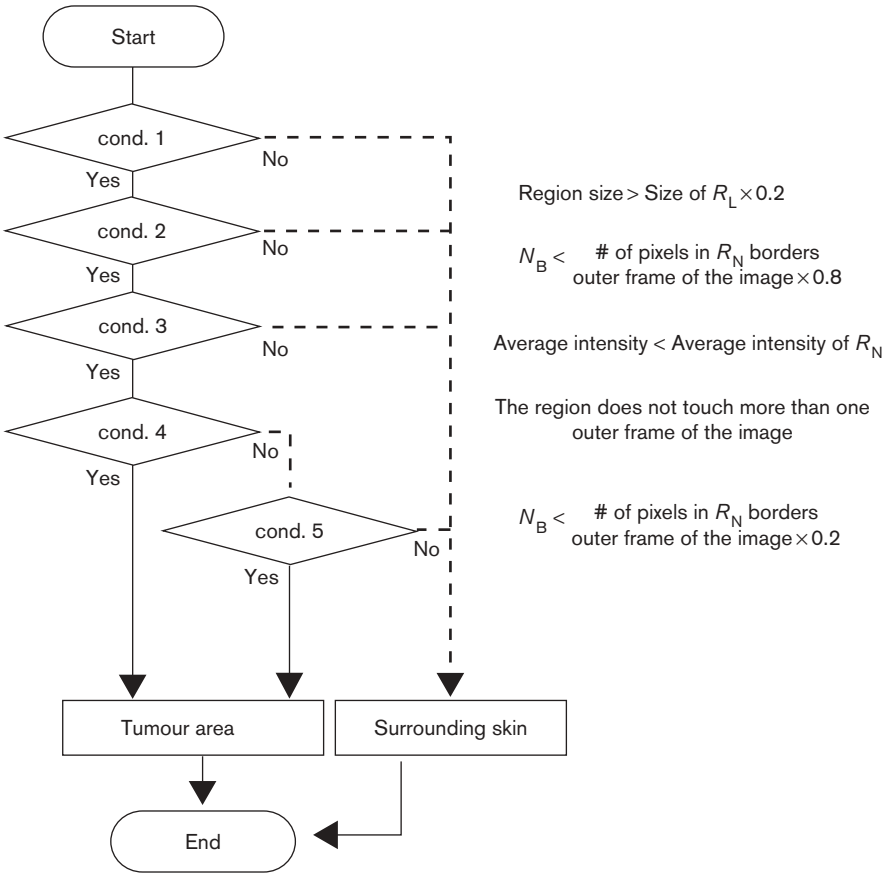
$$V_{in} \times \xi_{min} \leq V_{out} < V_{in} \times \xi_{max}$$

With the results of the preliminary experiments, we defined the following parameter values used in the method: $S = 7$, $\xi_{min} = 1.02$, $\xi_{max} = 1.07$.

Evaluation of the effectiveness of the coordinated tumour extraction algorithm

We investigated the influence of differences in tumour area extraction on the final diagnostic results using a linear classifier. Firstly, a total of 64 base parameters (62 parameters such as colour, texture, circularity and symmetry, used in our previous study [13], and the size of the tumour and the duration of the illness) were calculated from both conventional and coordinated extraction results. We then conducted an incremental stepwise input selection to select the effective parameters from the total of 64. During the repetitive process of incremental input selection, every linear classifier was evaluated with 'leave-one-out' cross-validation. This selection was continued until the classifier reached the maximum diagnostic accuracy.

Fig. 3



Flow chart of the tumour area selection phase of our tumour area extraction algorithm. Each segmented area is separated into tumour area and surrounding skin using this flow chart. N_B, number of pixel borders in the outer frame of the dermoscopy image; R_L, largest region; R_N, region bordering the external frame of the image with the longest boundary.

Results

Definition of the STA

The SDs of the average and major areas of tumour size were 8.9% and 7.4%, respectively, whereas the SD of non-medical individuals was 10.7%. The major area had a smaller divergence than the average area. We eventually selected the major area (selected by more than two dermatologists) as the STA.

Evaluation of extraction by non-medical individuals

The extraction by 10 non-medical individuals gave a precision of 97.0% and a recall of 90.2%.

Table 1 compares the SDs of the areas extracted by five dermatologists and 10 non-medical individuals by type of lesion. The SD of extraction by 10 non-medical individuals was 10.7% of the tumour size on average. The divergence was generally larger than that of dermatologists (8.9%), but the compositions were different.

Table 1 Standard deviation of extraction normalized[a] by tumour size (%)

	Tumour type			
	Clark naevus	Reed naevus	Melanoma	Average
Five dermatologists	9.2	8.9	8.2[b]	8.9
Ten non-medical individuals	11.3	9.2	10.3	10.7

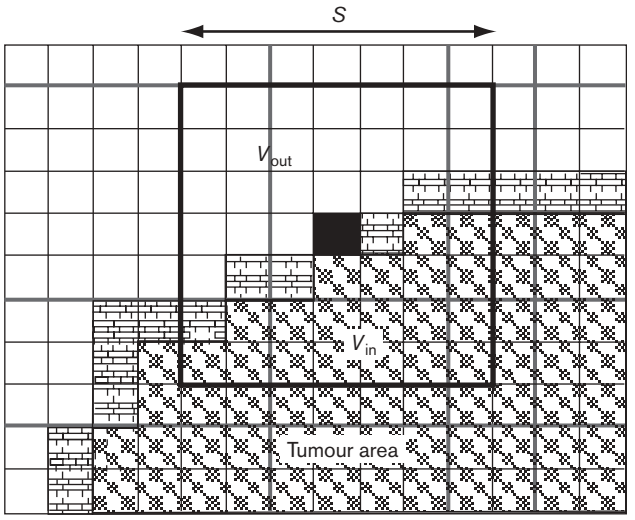
[a]Standard deviation was calculated image by image. Each value was normalized by the corresponding tumour size.
[b]Dermatologists selected the melanoma area with statistically smaller divergence than that of Clark naevus or Reed naevus ($P = 0.05$).

Dermatologists selected the melanoma region with a statistically smaller divergence than that of Clark naevus or Reed naevus ($P = 0.05$). In contrast, the area selected by non-medical individuals did not show this difference.

Evaluation of extraction by computer-based methods

The conventional computer-based method showed excellent precision (99.5%), but low recall (87.6%), because

Fig. 4



Concept of region-growing approach of our tumour area extraction algorithm. The temporarily decided tumour area was adjusted using a region-growing approach. For every border pixel (black pixel), the average intensities of the inside and outside of the tumour area, V_{in} and V_{out}, were calculated. In the corresponding equation, we regard the border to be ambiguous and treated the $S \times S$ area as the tumour area.

Table 2 Comparison of the tumour extraction results

	Precision (%)[a]	Recall (%)[b]
Ten non-medical individuals (average)	97.0	90.2
Conventional computer-based tumour extraction algorithm	99.5	87.6
Our coordinated computer-based tumour extraction algorithm	94.1	95.2

[a]Precision = true positive/(true positive + false positive) = (correctly extracted area)/(extracted area).
[b]Recall = true positive/(true positive + false negative) = (correctly extracted area)/(tumour area).
These are opposing criteria and good extraction requires high levels of both precision and recall.

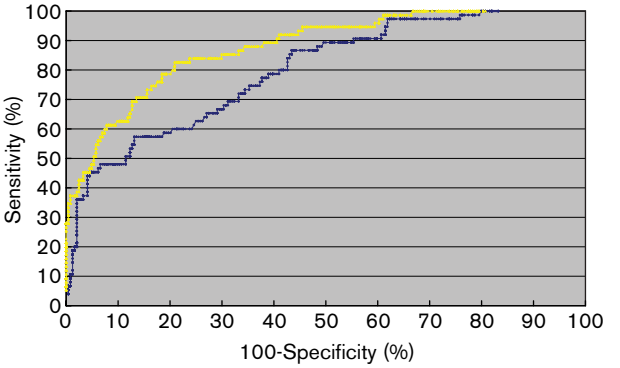
this method tended to extract the inner area of the STA. Our coordinated method additionally extracted 13.1% tumour area on the outer side of the conventional border and achieved superior results (precision, 94.1%; recall, 95.3%), see Table 2.

As our coordinated method did not give special consideration to the extraction of blue-white areas, namely regression or blue–whitish veil, each evaluation was different in images with these areas. The average extraction performance decreased (precision, 96.3%; recall, 87.7%) in nine dermoscopy images containing these areas.

Effectiveness of tumour extraction algorithm in terms of diagnostic accuracy

Nine input parameters were selected by the incremental stepwise input selection: asymmetry of red in the major axes, asymmetry of blue in the minor axes, minimum red value in the peripheral part, standard deviation of the green value in the peripheral part, proportion of areas with colour intensity equal to or less than 100 (maximum, 255), minimum blue and green values of the area occupying more than 0.05% of the total area, tumour duration and tumour size.

Fig. 5



Receiver operating curve of linear discriminant analysis based on conventional tumour area extraction results (blue line) and our coordinated extraction results (yellow line). Our coordinated tumour extraction algorithm improved the diagnostic accuracy relative to that of the conventional method. The area under the ROC increased from 0.795 to 0.875 with our coordinated tumour extraction algorithm. When the diagnostic threshold was defined at a sensitivity of 80%, our extraction method showed approximately 20% better accuracy in specificity.

Figure 5 compares the receiver operating curve (ROC) of the linear classifier based on conventional thresholding tumour extraction and coordinated extraction. The area under the ROC increased from 0.795 to 0.875. When the sensitivity was 80%, each linear classifier showed a specificity of 61.1% (conventional) or 81.1% (coordinated). Examining this result in detail, the specificity of Clark naevus improved from 60.1% to 77.1%, whereas that of Reed naevus improved from 57.1% to 85.7%.

Discussion

Our objective was a quantitative analysis of PSLs. As we did not want to handle exceptional cases at this time, we used certain limitations, described in the 'Materials' section, to select the dermoscopy images included in this study.

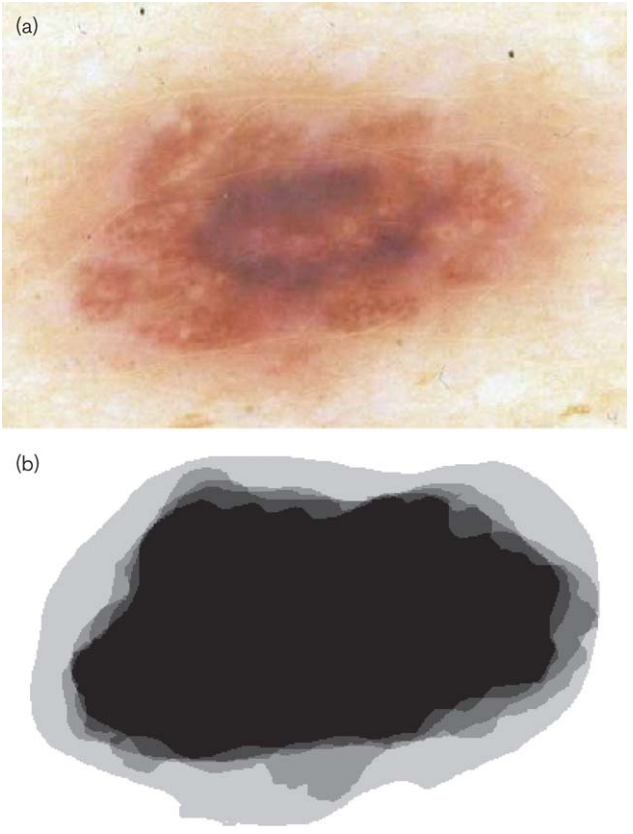
The conventional evaluation criterion for tumour area extraction, the 'extraction rate', was not sufficient for quantitative analysis because it evaluated only whether the extracted area was 'acceptable' or 'not acceptable' for each dermoscopy image. Therefore, we introduced new evaluation criteria: precision and recall [22]. These criteria are commonly used in engineering. Precision

indicates how accurate the extracted area is and recall indicates how well the tumour area is extracted. These are opposing criteria, and good extraction requires high levels of both precision and recall.

The evaluation of extraction demonstrated that there were substantial differences between dermatologists. In general, it seems reasonable to use the average area as the STA. However, the SD of the average area was large (8.9%), and it may contain inaccuracies. One of the dermatologists, not always the same physician, tended to extract larger areas than the others in many cases. Therefore, we defined the major area (SD = 7.4%), representing the area selected by more than two dermatologists, as the STA.

Figure 6 shows an example of extraction with large divergence. In this example, the SDs of the selected area were 17.1% and 10.5% of the tumour size for the average area and major area, respectively. Although this figure shows one typical example, the same trend was observed in many cases.

Fig. 6



An example of manual extraction in Clark naevus. (a) A dermoscopy image. (b) Results of manual extraction by five dermatologists. The standard deviation of the average area was 17.1% and that of the major area was 10.5%.

As the results of Schmid-Saugeon *et al.* [19] and Guillod *et al.* [21] showed a standard deviation of approximately 4–8%, we concluded that extraction by a dermatologist could not be regarded as the gold standard. Consequently, we set the target accuracy of automatic tumour area extraction as more than 95% [$\sim$100%–(8.9%/2)] for both precision and recall.

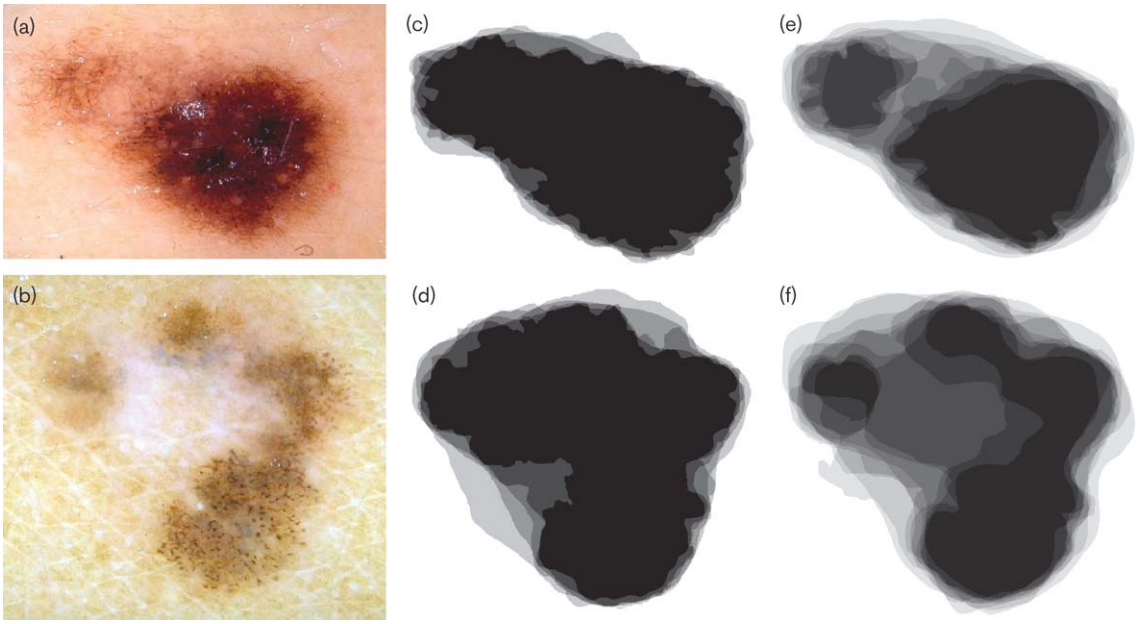
The average extraction by non-medical individuals showed a precision of 97.0% and a recall of 90.2%, indicating moderate differences in extraction compared with dermatologists. This result indicated that the area extracted by non-medical individuals was smaller than the tumour area (STA), and approximately 10% of the tumour area was not extracted.

The SD of the area extracted by non-medical individuals was generally larger than that of dermatologists, except for Reed naevus. Reed naevus tends to have a distinct border and therefore does not require such a specialist knowledge to decide the tumour area. The SD of the extracted area of melanoma by dermatologists was statistically smaller than that of other tumours, even though some melanomas have an ambiguous border. Dermatologists have a common knowledge about melanoma with regard to colour, shape, texture, and so on, and therefore they selected a similar area and the divergence was low. These results may also be partially due to the careful evaluation of malignant tumours. Because Clark naevi often have an ambiguous border, the SD of selection in this case was larger than that of the two other types of tumour.

Figure 7 shows examples of extraction with large differences between dermatologists and non-medical individuals. Dermatologists tended to recognize islands of lesions as one group. In contrast, non-medical individuals extracted each island separately. In addition, there were differences in the choice of regression or blue–white areas. Some non-medical individuals did not recognize the whitish areas as lesions, whereas others selected similar non-lesional areas. Large divergences between the two groups of examiners were generally observed in cases with an ambiguous border.

The extraction area of the conventional computer-based method showed a precision of 99.5% and a recall of 87.6%. The precision of this result was excellent, but low recall indicated that the extracted area was too small and that almost all of the extracted area was in the tumour area. Because it is necessary to extract an adequate tumour area when building an automatic melanoma diagnostic system, this method is not ideal. In particular, the characteristics of the peripheral parts of PSLs are important for the diagnosis of melanoma, so that inadequate extraction may result in the loss of important information. A precision of 100% can easily be achieved if

Fig. 7



Differences in extraction between dermatologists and non-medical individuals. (a) A dermoscopy image of a Clark naevus. (b) A dermoscopy image of melanoma *in situ*. (c, d) Extractions by five dermatologists. (e, f) Extractions by 10 non-medical individuals [(e) precision, 99.2%; recall, 78.0%; (f) precision, 94.7%; recall, 63.9%].

the extracted area is totally included in the tumour area. Other computer-based methods using the clustering technique showed a similar trend; they had high precision but low recall when compared with the results of dermatologists [20].

In the tumour area extraction task, the extracted borders selected by the dermatologists were shifted to lighter areas from the statistically optimum border calculated by the computer. The coordinated method used in this study considered this matter and achieved better results with high precision and recall. This indicates that our method achieved results close to those of the dermatologists.

The area under the ROC increased from 0.795 to 0.875 with our tumour area extraction algorithm. Thus our algorithm improved the diagnostic accuracy by 8.0% on average in this case. In particular, an improvement in the diagnostic accuracy of Reed naevus was apparent when the sensitivity was 80%. The mean discrimination capability between melanoma and Reed naevi was specifically improved with our tumour extraction algorithm.

In general, border and colour characteristics are important for the diagnosis of PSLs. Because the conventional tumour area extraction algorithm did not extract a sufficient tumour area, only the inner regions of melanomas with an ambiguous border were extracted, so that the original border properties were lost. In addition, Reed naevus often has similar colour characteristics to melanoma; therefore, distinguishing between Reed naevus and melanoma is difficult with incomplete tumour extraction. Our coordinated tumour extraction algorithm extracted the important peripheral part of the tumour and was useful in distinguishing between Reed naevi and melanomas with improved specificity. It should be noted that the linear discriminant analysis applied here is not for seeking diagnostic accuracy, as in other research, but for investigating the diagnostic impact of tumour extraction results using the same parameters.

Our algorithm did not contain special treatment for blue–whitish areas. The extraction of these areas often misidentifies reflections or other artefacts, because these regions show wide variations and cannot be specified definitely. However, these areas represent important features for the diagnosis of melanoma [1]. Our future plan is to focus on this problem and to develop an efficient algorithm and to conduct quantitative analysis. This coordinated tumour extraction method is already uploaded and is used on our web-based melanoma diagnostic system.

Conclusions

In this study, we investigated the quantitative evaluation of the manual extraction of tumour areas by dermatologists using dermoscopy images, and compared the results with those of non-medical individuals and two types of computer-based method. Differences in extraction between dermatologists and non-medical individuals were especially prominent in the choice of island-shaped, regressive or blue–white areas.

The conventional computer-based tumour area extraction algorithm did not extract sufficient tumour area. The coordinated computer-based method introduced the region-growing approach to bring the results closer to those of dermatologists. As a result, the method extracted the important peripheral part of the tumour and achieved almost equivalent results (precision, 94.1%; recall, 95.3%).

The area under the ROC increased from 0.795 to 0.875 with refinement of the tumour extraction algorithm. More importantly, the diagnostic accuracy improved by approximately 8% on average and 20% in specificity when the sensitivity was 80%.

We have confirmed quantitatively that the extraction of the peripheral part of the tumour is important for computer-based diagnosis.

References

1 Stolz W, Braun-Falco O, Bliek P, Kandthaler M, Burgdorf WHC, Congnetta AB. Color atlas of dermatoscopy. 2nd ed. Berlin: Blackwell Publishing; 2002.

2 Meyskens FLJr, Berdeaux DH, Parks B, Tong T, Loescher L, Moon TE. Cutaneous malignant melanoma (Arizona Cancer Center experience). I. Natural history and prognostic factors influencing survival in patients with stage I disease. *Cancer* 1988; **62**:1207–1214.

3 Buttner P, Garbe C, Bertz J, Birg G, d'Hoedt B, Drepper H, *et al.* Primary cutaneous melanoma. Optimized cutoff points of tumor thickness and importance of Clark's level for prognostic classification. *Cancer* 1995; **75**:2499–2506.

4 Garbe C, Buttner P, Bertz J, Berg G, d'Hoedt B, Drepper H, *et al.* Primary cutaneous melanoma. Identification of prognostic groups and estimation of individual prognosis for 5093 patients. *Cancer* 1995; **75**:2484–2491.

5 Balch CM, Murad TM, Soong SJ, Ingalls AL, Richards PC, Maddox WA. Tumor thickness as a guide to surgical management of clinical stage I melanoma patients. *Cancer* 1979; **43**:883–888.

6 Soyer HP, Smolle J, Kerl H, Stettner H. Early diagnosis of malignant melanoma by surface microscopy. *Lancet* 1987; **2**:803.

7 Ercal F, Chawla A, Stoecker WV, Lee HC, Moss RH. Neural network diagnosis of malignant melanoma from color images. *IEEE Trans Biomed Eng* 1994; **9**:837–845.

8 Seidenari S, Pellacani G, Pepe P. Digital videomicroscopy improves diagnostic accuracy for melanoma. *J Am Acad Dermatol* 1998; **39**: 175–181.

9 Elbaum M, Kopf AW, Rabinovitz HS, Langley RG, Mihm MC Jr. Automatic differentiation of melanoma from melanocytic nevi with multispectral digital dermoscopy: a feasibility study. *J Am Acad Dermatol* 2001; **44**:207–218.

10 Rubegni P, Burroni M, Cevenini G, Peorotti R, Dell'Eva G, Barbini P, *et al.* Digital dermoscopy analysis and artificial neural network for the differentiation of clinically atypical pigmented skin lesions: a retrospective study. *J Invest Dermatol* 2002; **119**:471–474.

11 Rubegni P, Cevenini G, Burroni M, Peotti R, Dell'Eva G, Sbano P, *et al.* Automated diagnosis of pigmented skin lesions. *Int J Cancer* 2002; **101**:576–580.

12 Blum A, Soyer HP, Garbe C, Kerl H, Rassner G, Hofman-Wellenhof R. The dermoscopic classification of atypical melanocytic naevi (Clark naevi) is useful to discriminate benign from malignant melanocytic lesions. *Br J Dermatol* 2003; **149**:1159–1164.

13 Oka H, Hashimoto M, Iyatomi H, Argenziano G, Soyer HP, Tanaka M. Internet-based program for automatic discrimination of dermoscopic images between melanomas and Clark naevi. *Br J Dermatol* 2004; **150**:1041.

14 Green A, Martin N, Pfitzner J, O'Rourke M, Knight N. Computer image analysis in the diagnosis of melanoma. *J Am Acad Dermatol* 1994; **31**: 958–964.

15 Jackowski LXM, Goshtasby A, Yu C, Roseman D, Bines S, Dhawan A, *et al.* Segmentation of skin cancer images, Image vis. *Computing* 1999; **17**: 65–74.

16 Hance GA, Umbaugh SE, Moss RH, Stoecker WV. Unsupervised color image segmentation with application to skin tumor borders. *IEEE Eng Med Biol* 1996; **15**:104–111.

17 Schmid P. Segmentation of digitalized dermatoscopic images by two-dimensional color clustering. *IEEE Trans Med Imag* 1999; **18**:164–172.

18 Ganster H, Pinz A, Roehrer R, Wilding E, Binder M, Kitter H. Automated melanoma recognition. *IEEE Trans Med Imag* 2001; **20**:233–239.

19 Schmid-Saugeon P, Guillod J, Thiran JP. Towards a computer-aided diagnosis system for pigmented skin lesions. *Comput Med Imaging Graph* 2003; **27**:65–78.

20 Iyatomi H, Oka H, Miyake A, Kimoto M, Tanaka M, Ogawa K. A practical internet-based melanoma diagnostic system. In: IEEE, editor. Proceedings of IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), November 14–15, San Diego, CA, USA, 2005. Piscataway, NJ: IEEE Press; 2005. pp. 443–446.

21 Guillod J, Schmid-Saugeon P, Guggisberg D, Cerottini JP, Braun R, Krischer JH, *et al.* Validation of segmentation techniques for digital dermoscopy. *Skin Res Technol* 2002; **8**:240–249.

22 Rosenfeld L, Morville P. Information architecture for the world wide web. 2nd ed. San Francisco: O'Reilly; 2002.

23 Otsu N. An automatic threshold selection method based on discriminant and least square criteria. *Trans IEICE* 1988; **63**:349–356.