



MASTER OF ENGINEERING  
IN  
TELECOMMUNICATIONS SYSTEMS

**Final Report for Skin Lesion Project**

Faivre Matthieu 54147158

Year 2004/2005

Supervised by Professor P. Whelan

## **I/ Acknowledgements**

First, I would like to thank my supervisor Pr. Paul Whelan for his guidance in this project.

Thanks are also due to Mr Robert Sadleir and Mr John Mallon for supporting this work and giving me some clues to start this enterprise.

## **II/ Declaration**

I hereby declare that, except where otherwise indicated, this document is entirely my own work and has not been submitted in whole or in part to any other university.

Signed: .....

Date:

## Table of Content

I/ Acknowledgements .....	2
II/ Declaration .....	2
Table of content	
III/ Abstract .....	4
IV/ Introduction .....	4
V/ Specifications .....	4
VI/ Chapter 1 : Technical background .....	5
VII / Chapter 2 : Outline .....	6
1) Explanation of the selected parameters .....	6
2) Explanation of the NeatVision file .....	7
a) the segmentation .....	7
b) the shape features .....	10
c) the radiometric features .....	10
VIII/ Chapter 3 : Results and discussion .....	11
1) Description of the sample input images .....	12
2) Mathematical results .....	13
3) Appearance features .....	15
IX/ Chapter 4 : Problems raised .....	27
X/ Chapter 5 : Possible improvements .....	28
XI/ Conclusion .....	28
XII/ References .....	29
Global View of the NeatVision file .....	30

### **III/ Abstract**

“The aim of this project is to develop a medical imaging system in the area of skin feature analysis and classification with a view to helping the early medical diagnosis of skin melanoma. Early diagnosis of melanoma is currently essential for successful treatment, research shows that accurate early diagnosis by general medical practitioners is low. The automatic system proposed will be designed to support practitioners in the early diagnosis of skin melanoma and to improve the early clinical diagnosis success rate. This project will build on the VSG’s existing research program in colour texture analysis.”

### **IV/ Introduction**

Malignant melanoma is nowadays one of the main cancers among light-skinned populations around the world. The increase in ultraviolet radiation, due to holes in the ozone layer and the lifestyle of people, cause a huge increase in the number of melanomas diagnosed. The raise of this phenomenon was first noticed during the 30’s in the USA, where one person out of 100 000 suffered from skin cancer (per year). This rate increased in the middle of the 80’s to 6 per 100 000 and to 13 per 100 000 in 1991. These statistics are similar to the incidence rates observed in Europe.

Investigations have shown that the curability of skin cancer is close to 100% if it is recognized early enough and treated surgically. The mortality rate caused by melanomas in the early 60’s was about 70% and now a survival rate of about 70% is achieved, which is mainly due to the early recognition.

Melanoma is a malignant tumor of the melanin forming cells, the melanocytes. When a melanoma invades these melanocytes descend from the dermo-epidermal junction into the dermis. The depth of this invasion, or thickness, is considered as the single most important prognostic factor. Some scientists say that the relationship between thickness and survival is progressive but not continuous especially at the edges : under 0.76 mm thick, more than 95% of the tumors are curable, over 5 to 6 mm thick, most tumors are lethal.

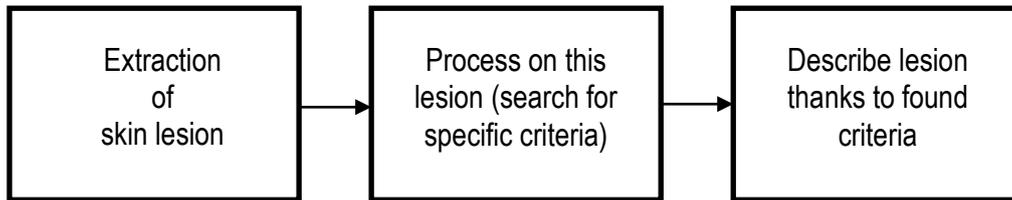
### **V/ Specification**

I have to design a system capable of helping dermatologists to diagnosis skin melanoma.

The system should highlight some features of human skin pictures (such as area, perimeter, compactness, circularity, mean of colour pixels, entropy, areas with mostly bluish or reddish pixels in the lesion) in order to make easier the diagnosis of doctors. This system should thus not be designed to give an opinion, just to bring out features of the image.

## VI/ Chapter 1 : Technical background :

To carry out this project, I split the work in 3 mains stages :



The extraction of the skin lesion is achieved by the segmentation.

The process to search for specific criteria takes different forms, as it is calculating mathematical parameters colour and shape features.

The description of the lesion is achieved by either the superposition of masks on the input image, or a specific colour process, or even a study of the mathematical values found.

In order to do this project, I read the papers supplied in the Student Cancer Pack and I mainly used 5 paper :

-***“Automated Melanoma Recognition”***, by H. Ganster, A. Pinz, R. Röhner, E. Wildling, M. Binder and H. Kittler.

I used this paper for the segmentation methods submitted. They are describing powerful methods used by colleagues, such as global thresholding, dynamic thresholding and 3D colour clustering. Actually they are using all these 3 methods and merge results (with some rules) to get their final segmentation. The features they focus on are : area, perimeter, polar distance, bounding box and the distance from the centroid (of the blob) to the bounding box. For the shape features, they calculated the roundness, compactness, extent and they divided the picture by 2 perpendicular axis. They switched to the HIS space and calculated minimum, maximum, mean, variance for hue and intensity components.

-***“Initial results of Automated Melanoma Recognition”*** by H. Ganster, M. Gelautz and A. Pinz, from the Institute for Computer Graphics, Technical University Graz, Austria.

This paper was majority interesting for the shapes features description. In this paper, the authors present a way of extracting, processing and labeling melanoma. To extract the feature of the skin lesion (the shape) from the background, they used a global thresholding in a bimodal histogram.

To find the threshold value, they use the HSV space (for hue, saturation and value) and look for local minima. This step was for the segmentation.

For the features to extract from the pictures, they worked on the area, perimeter and compactness of the blob (describing the lesion), they used a bounding box as well. For the colours, they processed the minimum, maximum, mean and variance of each colour channels of RGB space.

-***“Towards a computer-aided diagnosis system for pigmented skin lesions”*** , by P. Schmid, Joël Guillod, Jean-Philippe Thiran from the Swiss Institute of Technology.

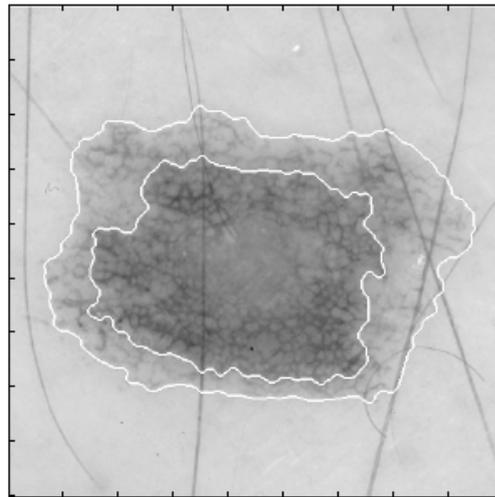
Best points of this paper are the description of the ABCD rules and the method to detect the skin boundary. They explain a way of removing hairs, inspired by Dullrazor. Nevertheless they stressed that the best position would be to have as less hairs as possible because of interferences of hairs in the segmentation process (they even talk about shaving the area before taking the picture).

For the segmentation they used an orientation-sensitive fuzzy c-means clustering that works on a LUV space. The transform (in a 2 dimension histogram) is called Karhunen-Loève. The boundary detection is based on colour differences and on morphological flooding. The main feature they are looking for is the symmetry of colours and textures.

-“**Colour segmentation for the analysis of pigmented skin lesions**”, by P. Schmid and S. Fisher from the Swiss Federal Institute of Technology.

In this paper the authors are explaining several methods to implement segmentation, based on filtering (low-pass, median), histogram analysis and clustering.

A result for a good segmentation is given :



This segmentation was obtained with an histogram information analysis and a fuzzy c-means clustering technique.

-“**Segmentation of Skin Cancer Images – a report**”, by Padmapriya Nammalvar from the DCU.

This paper is explaining clearly several methods such as adaptive smoothing and k-mean clustering.

## VII/ Chapter 2 : Outline

### 1) Explanation of the selected parameters :

The features I decided to highlight are the ABCD rules, which represent the basis for a dermatologist's diagnosis, that's why I decided to implement these rules for the description of the segmented skin lesions :

-A stands for asymmetry.

We can highlight this feature thanks to the circularity index (defined as the ratio  $\text{area}/\text{perimeter}^2$ ), compactness ( $\text{area}/\text{perimeter}$ ), position of the centroid within the lesion, a circle with the centroid as origin, a box bounding the lesion and eventually perpendicular axes centred on centroid.

-B for border irregularity.

The sharpness of transition from lesion interior to skin is a good sign of malignant lesion. The different (more or less) concentric layers of colours in a lesion can be brought out too, as these layers are sometimes hard to distinguish, even for a dermatologist.

-C is for colour variation.

Here the idea is mainly to tell the user of the program where “dangerous” shades of colours are located, especially blue-grey or bluish shades. Reddish tones can be highlighted as well. The “quantification” of colours, for example the mean, the highest and lowest level of a colour or the entropy can be helpful for the dermatologist.

-D stands for differential structures or, for other scientists for diameter greater than 6 mm.

For this characteristic we can calculate the diameter of the bounding circle. The use of the area, perimeter, standard deviation or a switch to the HIS domain can be relevant for this study.

Protrusions and indentations are not expressly quoted, but there are a good clue of the regularity of a lesion and the way it is spreading.

## 2) Explanation of the NeatVision file :

### a. the Segmentation

So for this step I choose to implement a non-destructive method, especially for the hairs and glints of light. I didn't want to remove possible hairs as they were not present in the majority of the input pictures. Furthermore, if I would have suppressed them and then applied an interpolation to limit destructive effects of hairs-removing, I would undoubtedly have lost data (or at least modified the original image) and the result is not satisfying as we still can guess where hairs were. The problem of light glint, which is present on several pictures, was tackled with the same idea than hairs over-segmentation.

All the previous papers are describing one or several ways of segmenting and describing skin lesions and I decided to build the segmentation around a fuzzy k-mean clustering, helped by a blurring pre-processing step.

The main idea of the segmentation was, instead of analyzing where potential problems were (hairs or glint) and suppressing them, to “dilute” them with the rest of the data, in order to minimize their effects on the segmentation. So I decided to blur the input picture and as a result I got a better distinction of the most important shape in the image, which is the skin lesion in its whole (and not only a part of it, due to hairs corrupting the segmentation).

The blurring was applied on a grey-scale image, so the input colour image has to be converted before. A Rectangular average filter was used for the blurring process, this is a kind of convolution.

Nevertheless this mathematical transformation inserts a black area situated along the edge of the picture, as the convolution needs a certain amount of pixels from the image to process one pixel on the output image, resulting in a black frame. This frame is not so a huge nuisance, as almost all pictures are well centred. So it means that the pixels lost (on the frame) because not processed by the convolution, have less importance than the ones, constituent the hairs, that could have been removed with a hair-removing system, especially for those situated at the border of or crossing all the lesion.

Firstly, I wanted to isolate hairs and to subtract them from the input image, I achieved to do it, detection of hairs was good but, once subtracted, the result was not sharp enough to my opinion as they were gaps (sometimes huge) between same parts of the lesion (see 27). These gaps, even processed with another convolution (with an appropriate input array, in order to do a kind of interpolation), were still interfering in the segmentation by “over-chopping” the skin lesion, resulting in many skin lesions instead of only one.

The segmentation, properly speaking, is achieved with a Grey scale cluster. The k-means Clustering is the process of grouping finite set of objects into subsets. The objects are organised into an efficient representation that describe the population being sampled. This study uses the unsupervised clustering technique based on the k-means algorithm, which is the simplest and most popular algorithm between all the iterative clustering algorithms.

The number of clusters represents the input parameter for the k-means algorithm and the value of this parameter is dependent on the complexity of the image. To determine the optimal number of cluster could be a very difficult task as this parameter is image dependent.

I set the number of clusters to 3; in other words it means that the input image will be described with only 3 grey values, which is sufficient to detect our main blob.

After the Grey scale cluster, there is a threshold with a really low input value to get back the biggest blob in the image (which has the lowest grey level).

Then the mask operation is suppressing the frame created by the Rectangular average filter.

After these steps the segmentation has to be improved with a block isolating the biggest blob in the image. Once it is done, this blob is filled with white colour to get a plain white blob instead of a shape full of “swiss cheese” holes.

The Erode and Dilate blocks are here only to definitely suppress the effects of hairs. The Erosion will erode or “compress” the white areas whereas the Dilation will expand these areas in order to find again a blob as similar as possible as the lesion.

Actually this succession of Erode and Dilate blocks are here only for one picture : SkinLesion\_Orig3, for all the others pictures these operations are useless in the sense that we already have a good match for the lesion detection .

At this point of the file and after an ANDing of the input image with the blob obtained, we have the skin lesion isolated from the rest of the input image. We can now process it to calculate radiometric and shape descriptors.

Here are a few binary shapes for the lesion detection :

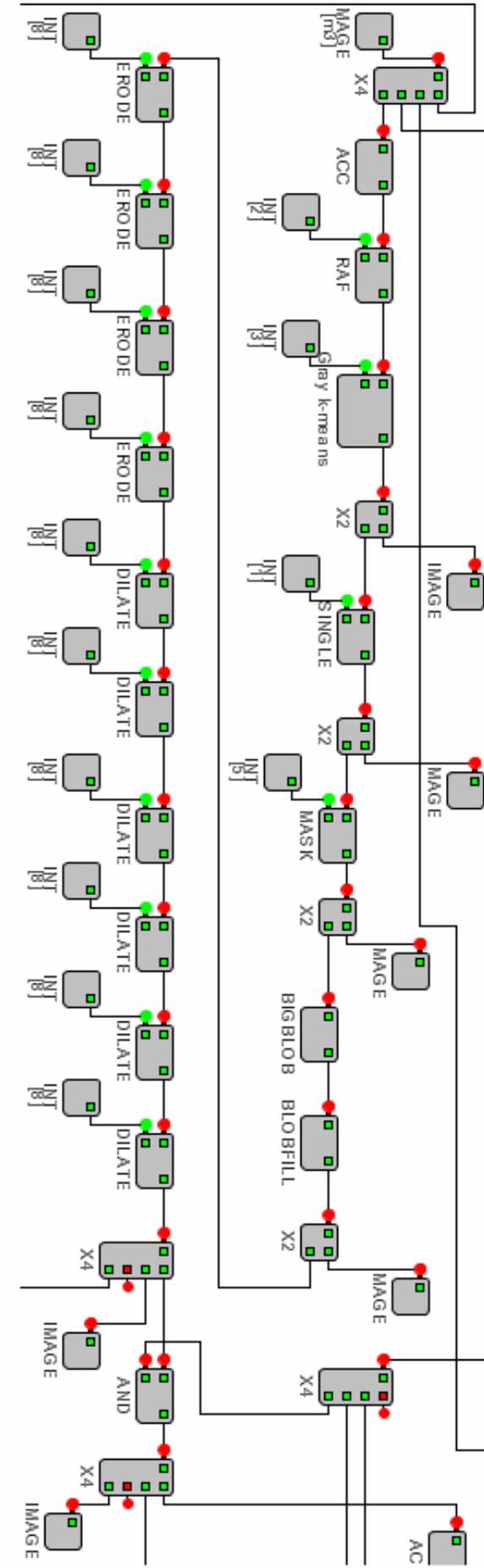


M4



M1

Final report for skin lesion project



A screen-shot of the segmentation part of the whole file. Start of the development is at the top.

## b. Shape features :

**The area** : to get this parameter, I simply used a White pixel counter with the binary image of the blob obtained after segmentation.

**The perimeter** : before counting the white pixel, I applied a Binary Border detector which gives as a result a one-pixel-width line describing the bend of the lesion.

**The circularity index** : is defined as the ratio  $\text{area}/(\text{perimeter})^2$ , so thanks to some mathematical blocks in NeatVision, it was possible to get this parameter. It gives an indication of the degree of “roundness” of the blob.

**The compactness index** :  $\text{area}/\text{perimeter}$ . Give an idea of the “shrinkness” of the blob, if it spreads a lot or not.

**The centroid** is giving the centre of gravity of the blob.

**The standard deviation** : is a measure of the spread of a set of values from the mean value. It is the square root of the variance and this is a measure of the dispersion.

**The bounding box** : is actually the box including or containing all the white blob.

**The gravity circle** : it is drawn with a centre equal to the centroid of the blob and a radius equal to the distance between the first corner of the bounding box and the centroid , this is giving an idea of the shape of the blob and its distribution in a circle (and in a box).

**The perpendicular axis** : centred on the centroid (x,y). To draw these axis I took as coordinate (0,y) , (width of picture, y) , (x,0) , (x, height of picture).

**The diameter** : in pixels, it is calculated from the radius of the gravity circle.

## c. The radiometric features :

**The colour analysis of the blob** : to get the most important colour shades of the lesion and their distribution, I used an Adaptive smoothing block.

The adaptive smoothing is a nonlinear smoothing which adapts pixel intensities to the local attributes of an image on the basis of discontinuity measures. Nonlinear smoothing preserves important features and also removes noise. A feature used is the local and contextual discontinuity. To measure those discontinuities, 4 detectors along the four directions are defined. The block has some inputs : the picture to process, the number of iterations (1 to 50), the variance (0.1 to 0.9) and the diffusion parameter (1 to 20).

The idea is to find the best suitable inputs parameters for all the pictures we want to work on.

The goal of adaptive smoothing is here similar than convolution but is more efficient as we can better adjust different parameters to emphasize specific characteristics.

When hairs, glint and small interfering structures have been smoothed and “threadbare”, we can try to describe this “trikling” picture with a few grey levels, thanks to a colour to grey scale cluster.

Once our lesion is a 4 grey level picture, we have to isolate each grey level, to remove eventually tiny black pixels (surrounded by white areas) and then to apply a binary border to get only the shape of the area.

Finally I added these results to the input image to get a better view of the result.

**The bluish, grey-bluish and reddish areas** : to find the places where the amount of these particular colours are situated, I switch to the RGB domain and work on the red and blue component, thanks to double thresholds that are framing the values of red and blue.

To finish I removed isolated white pixels and processed to dilation and erosion to make the shapes round. Another good indicator that is built-in with NeatVision is the Histogram view, that is helpful to “weigh” results from bluish and reddish pixel detection.

**The mean of the grey scale pictures** : it gives an idea of the largest grey level pixels in the picture, so we can deduce, among other things, if the global picture is light or dark.

**The entropy** : is a measure of the disorder of a system.

**The variance** : is calculated by taking the sum of the differences between the observed value and the series mean value, and dividing by the sample size. This is a clue for dispersion.

The **lowest and highest grey level** of the picture are calculated as well.

All the last 5 parameters are helping to characterize the picture and its repartition.

These parameters are extracted from hue and intensity components of a HSI domain too.

## VIII/ Chapter 3 : Results and discussion

I decided, for a more understandable reading, to split all the shape and radiometric features into 2 new groups :

- **Mathematical values** : area, perimeter, circularity index, compactness and standard deviation of the lesion blob then mean, variance, entropy, highest and lowest value of the grey scale image and of the hue and intensity component of the input image.

All the results of this group are gathered in a tab.

-**Appearance features** : centroid, bounding box, gravity circle, perpendicular axis, colour shades, bluish and reddish areas of the lesion.

I will first quickly describe the input images with a “first look” description of each image.

**1) Description of the 9 given sample images :**



**H1:**  
Vague and diffused shape, red and brown-greyish hues, visible pixels



**M1 :**  
Dark brown spot, some thin light hairs, a small dense and compact crater at 11, some glint at the top



**M2 :**  
Similar than M1, regular shape, some glint on the right, no hairs



**M3 :**  
Elongated spot, top seems to be a crater, bottom a rise with greyish and brown colours, a few glint on top



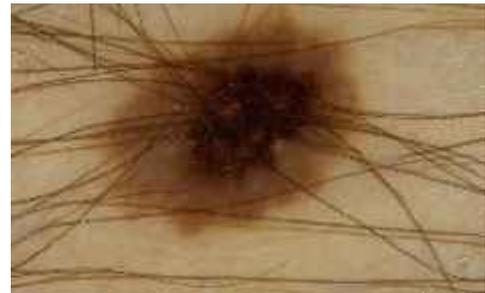
**M4 :**  
Irregular shape, spreading like magma, glint on the middle, 2 main colours light and extreme dark brown



**SkinLesion\_orig\_1 :**  
Egg-shaped blob, lot of greyish and reddish areas, lot of dark hairs around and over lesion, glint at the border, 1 main dark spot on top



**SkinLesion\_orig\_2 :**  
A massive brown and greyish spot, large hairs glint at the center near the blue-greyish area



**SkinLesion\_orig\_3 :**  
Round-elongated blob with 2 main visible colours: light brown at border and dark at the center, lot of thick hairs sometimes mingled together



**SkinLesion\_orig\_4 :**  
A diffused light brown lesion, with a tiny brown spot at the center

**2) Mathematical results :**

Now I will present the mathematical values and analyse how we obtained them:

	<b>Circu.</b>	<b>Compac</b>	<b>Area</b>	<b>Perim.</b>	<b>Mean</b>	<b>Entropy</b>	<b>High</b>	<b>Low</b>	<b>Var.</b>	<b>Diam.</b>	<b>SD</b>
H1	0.049	19.9	8077	404	125	2.22	214	48	774	120	27.83
M1	0.088	30.15	10344	343	115	3.88	255	0	2155	130	46.43
M2	0.088	20.73	4871	235	98	3.86	255	21	1667	90	40.83
M3	0.055	20.44	7564	370	110	3.98	255	16	1248	140	35.33
M4	0.059	41	28211	688	92	3.42	200	24	690	240	26.26
Skin 1	0.062	29.82	14433	484	142	5.19	253	1	2120	142	46
Skin 2	0.053	38.5	27713	719	92	5.05	253	4	1816	236	42.6
Skin 3	0.068	27.5	11113	404	111	4.74	184	5	1549	134	39.36
Skin 4	0.042	25.5	15592	611	133	4.34	248	70	374	232	19.3
<b>Mean column</b>	<b>0.63</b>	<b>28.17</b>	<b>11135</b>	<b>473</b>	<b>113</b>	<b>4.08</b>	<b>235</b>	<b>21</b>	<b>1377</b>	<b>162</b>	<b>36</b>

Circu.= circularity index| Compac= compactness index |Var.= variance| Diam.= diameter| SD= standard deviation

For this table, I would like to say that area and perimeter are perhaps not calculated the same way than it is in several papers (especially [3] ) as I took the number of white pixels (in the binary mask) for the area and the number of white pixels forming the one-pixel-width line for the perimeter. I don' t think that Neat Vision is taking care of diagonals (particularly for the perimeter) between two 8-connected pixels. The area is just the number of white pixels, which is close but not the real surface, as the real area is the number of white pixels times the surface of each pixels, so there is an uncertainty for the perimeter and the area but this uncertainty is affecting every picture.

Nevertheless these parameters are still good clues because they are representative of a trend, especially the ratio between maximum and minimum of each index and compared to the mean.

In the [2] paper there is no given unit for the values so we can only compared the ratio delivered with the values of malignant, dysplastic and benign lesions.

The ratios of extreme values for compactness, area, perimeter and variance are quite similar between this table and the one in [3], so at least we have a representative panel of pictures..

Circularity index is giving clue about how round a lesion is, so naturally the maximum values are reached for M1, M2 and Skin3, which are rather round.

The compactness index is therefore supplying data about how dense is a spot, or if it is spreading irregularly and, once again, the maximums are reached for the 2 most irregular spots : M4 and Skin2.

The mean, high and low grey-scale values are worked out to give an idea of the colours band of an image and to situate it in comparison to the rest of data or the average values.

Variance and standard deviation are indicative of dispersion and are mathematically really close, as standard deviation is the square root of the variance. The variance is the sum of the difference between the observed value and the series mean value, and dividing by the sample size. For example,

**Final report for skin lesion project**

Skin4 is very light and homogeneous, no sharp transition and logically it has the smallest variance and standard deviation.

Diameter, as well as area and perimeter, is given in pixels ; circularity and compactness have no unit ; mean, high and low are given in a range [0;255] for a grey scale value on 8 bits.

The user has to compare these values with what he knows or what he already has in memory ; for example he could analyse table of data processed from this NeatVision file with what he already processed and saved in a database, so the more tests he does, the closer from reality the means will be.

All these radiometric features have to be weighted because, for example, M4 is the biggest blob, so has the biggest area and perimeter, but it is the biggest picture as well, perhaps with a standardized picture acquisition system M4 would have appeared less enormous...

Here is a table of values for, on the left, the Hue component and on the right the Intensity component :

	<b>Mean</b>	<b>Entropy</b>	<b>High</b>	<b>Low</b>	<b>Variance</b>		<b>Mean</b>	<b>High</b>	<b>Low</b>	<b>Variance</b>	<b>Entropy</b>
	63	2.32	250	5	8603		126	214	48	774	2.22
M1	23	3.05	255	4	312		116	255	0	2155	3.88
M2	23	3	255	6	551		98	255	21	1667	3.85
M3	34	2.63	255	2	6175		110	255	16	1248	3.98
M4	54	2.61	255	5	7765		92	200	24	689	3.42
Skin 1	77	3.60	255	0	11418		142	253	1	2120	5.19
Skin 2	22	3.26	255	0	1773		92	253	4	1816	5.05
Skin 3	22	2.70	255	0	48		111	184	5	1549	4.74
Skin 4	13	2.48	25	1	10		133	248	70	374	4.35
<b>Mean Column</b>	<b>36.78</b>	<b>2.85</b>	<b>228</b>	<b>2.5</b>	<b>4072</b>		<b>113</b>	<b>235</b>	<b>21</b>	<b>1377</b>	<b>4.08</b>

I want to point that the 9 pictures have different formats (.gif and .bmp), different sizes and thus different resolutions. It means they are not “homogeneous”, so the system will have to be polyvalent or non-specialized.

The lighting, angle spot and distance between lesion and lens are really disparate too. So for the Intensity component all the values are similar to the ones of the grey-scaled input image, perhaps the colour\_to\_HSI block is using a greyscale for the intensity component.

For the hue component, results are quite different and help to complement each other.

Entropy is a measure of disorder of an image so the bigger, the “messier”. Mess is the presence of pixels with different values in the same area, with no apparent frame.

We can notice that light-coloured images with “smooth” transition (in other words with edges hard to guess between to different shades) have a small entropy (for example M1 and Skin4).

Output images from the Hue component are really different from the ones of the intensity component, for example all the pictures are very dark (sometimes flecked or with small white areas), so it is normal that results are different than the ones for grey-scaled images.

The hue component is looking for similar shades and most of the same light shades appear dark and the opposite for dark shades.

To have a far better view of the hue component, the Manipulate menu is really useful and more especially the Probe number 4, which gives you a good lead of the nuance tendency of the input images.

Hairs and glint have an incidence on the entropy, for example Skin1 and Skin3.

The panel of sample images is, although not too dense, well representative of what we can watch in the literature and this is confirmed with the results I got, which are similar and have the trend to those I found in the supplied materials [3].

I am naturally not allowed to dissect these values and to give my personal opinion about the clues highlighted with these results but I can explain how I got them and in which manner we can use and trust them for a better diagnosis.

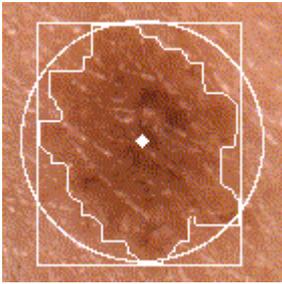
So all these parameters have to be studied by a specialist, compared together and, in addition, extreme, mean, highest and lowest values of these features can be compared with equivalent values of other specialists databases.

### **3) Appearance features :**

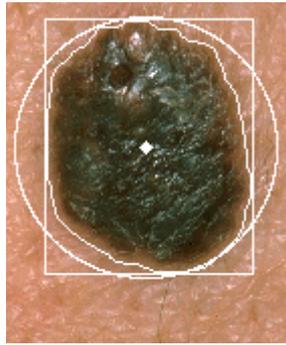
I would like to stress that I didn't present all the possible output images, but only the ones that seem clear and interesting to me. Nevertheless the user can put up a lot of different pictures at different stages of the process, so that he can have access to what he is looking for ( for example a black and white image of the lesion and its bounding box instead of a superposed image of the box on the original photo ; it depends of what the user most like and what he wants to watch ).

Now I will present the Appearance features, which are :

a. The centroid, gravity circle and the bounding box on the original image :



H1



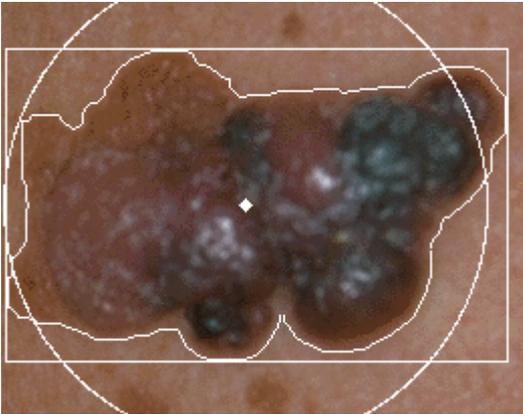
M1



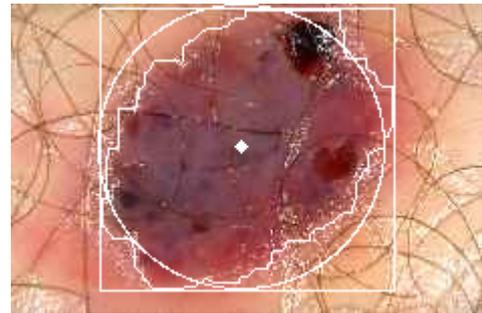
M2



M3



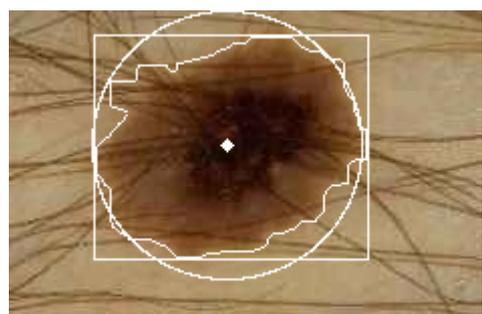
M4



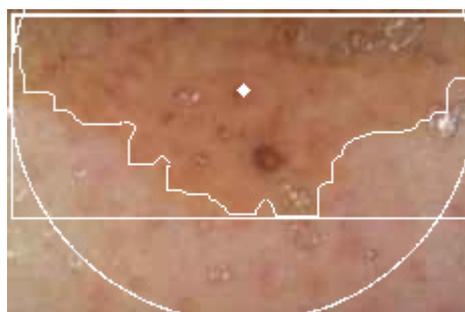
SkinLesion\_orig\_1



SkinLesion\_orig\_2



SkinLesion\_orig\_3



SkinLesion\_orig\_4

All these pictures are presenting the input image with a mask made up of the border of the lesion, the bounding box, which is the smallest that includes the all area of the blob, the centroid and the gravity circle, centred on the centroid.

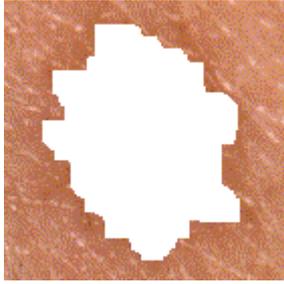
This superposition provides a view for the spatial arrangement of each lesion and shows how the spot fits in elementary forms.

Furthermore, the displaying of the centroid allows the user to see the exact centre of the shape, without being disturbed with "way out" colours blobs (Skin2) or elongated shape (M4) or even difficulty to distinguish the beauty spot (Skin4).

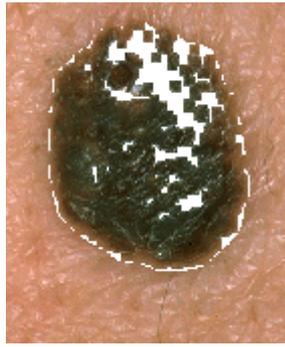
This representation is helpful to split into the main textural or colour features composing the lesion, for example M3, with a kind a quite light crater at the top and brown-greyish textures at the bottom.

The circle and the lesion border make easier the detection of small dark blobs at the periphery. All these parameters are reliable as they are directly extracted from mathematical values describing the blob.

b. The location of blue grey-bluish areas :



H1



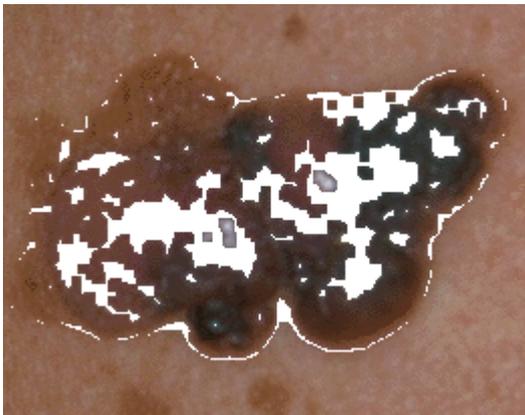
M1



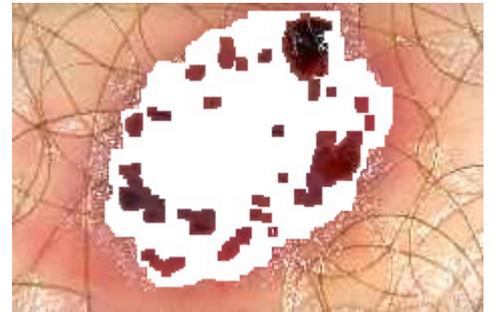
M2



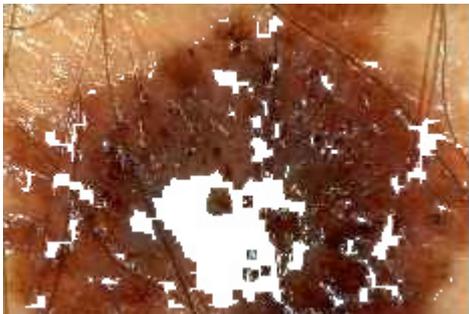
M3



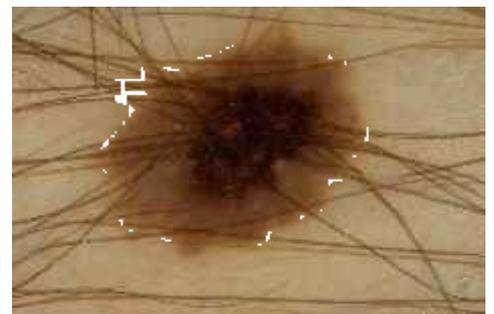
M4



SkinLesion\_orig\_1



SkinLesion\_orig\_2



SkinLesion\_orig\_3



SkinLesion\_orig\_4

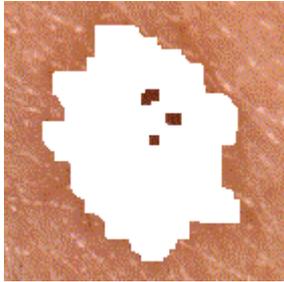
These pictures come from the red component of a RGB signal, “thresholded” with 2 values (127 and 65) that are supposed to “filter” and keep the blue or bluish pixels.

Actually this is working quite properly, even if the result could seem corrupt, especially for H1 and Skin4, for which it is hard to trust that there is so many blue pixels, but, in fact if you zoom on these 2 pictures you will see a kind of framework composed with light purple or violet pixels, so they are not visible at the first sight, but they are present in the picture all the same and that’s why they are highlighted.

For M1 and M2, there is really few blue pixels, but glint is corrupting the detection by adding touch of clear colours, resulting in mix of shades.

Globally results are satisfying, for example Skin1, which is a sea of blue shade pixels or for Skin2, on which we can perfectly determine where is the blue-greyish spot.

c. The location of reddish areas :



H1



M1



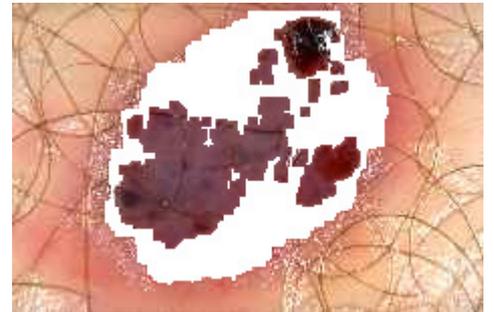
M2



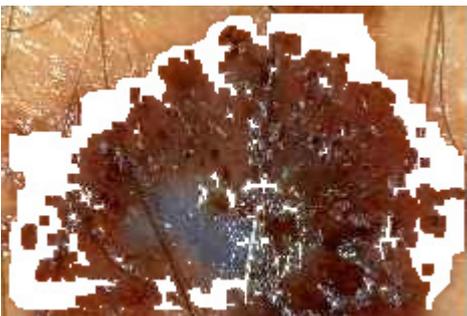
M3



M4



SkinLesion\_orig\_1



SkinLesion\_orig\_2



SkinLesion\_orig\_3



SkinLesion\_orig\_4

Here the idea is the same than for the detection of bluish areas, except the parameters for the double threshold, which are 202 and 130.

Once again, the fact that for 2 of the sample pictures (H1 and Skin4) the result is showing so much red or reddish pixels can be understand as, after zooming, you will see a matrix of flesh-coloured pixels, which has some red components.

The detection of red (and blue) areas is based on the colours inside the lesion, so there is no study of the pixels forming the skin around and outside the lesion.

For M1, M2 and M4, it is obvious that there is hardly no red pixels.

For M3, the crater situated in the top half is clearly highlighted for having red pixels.

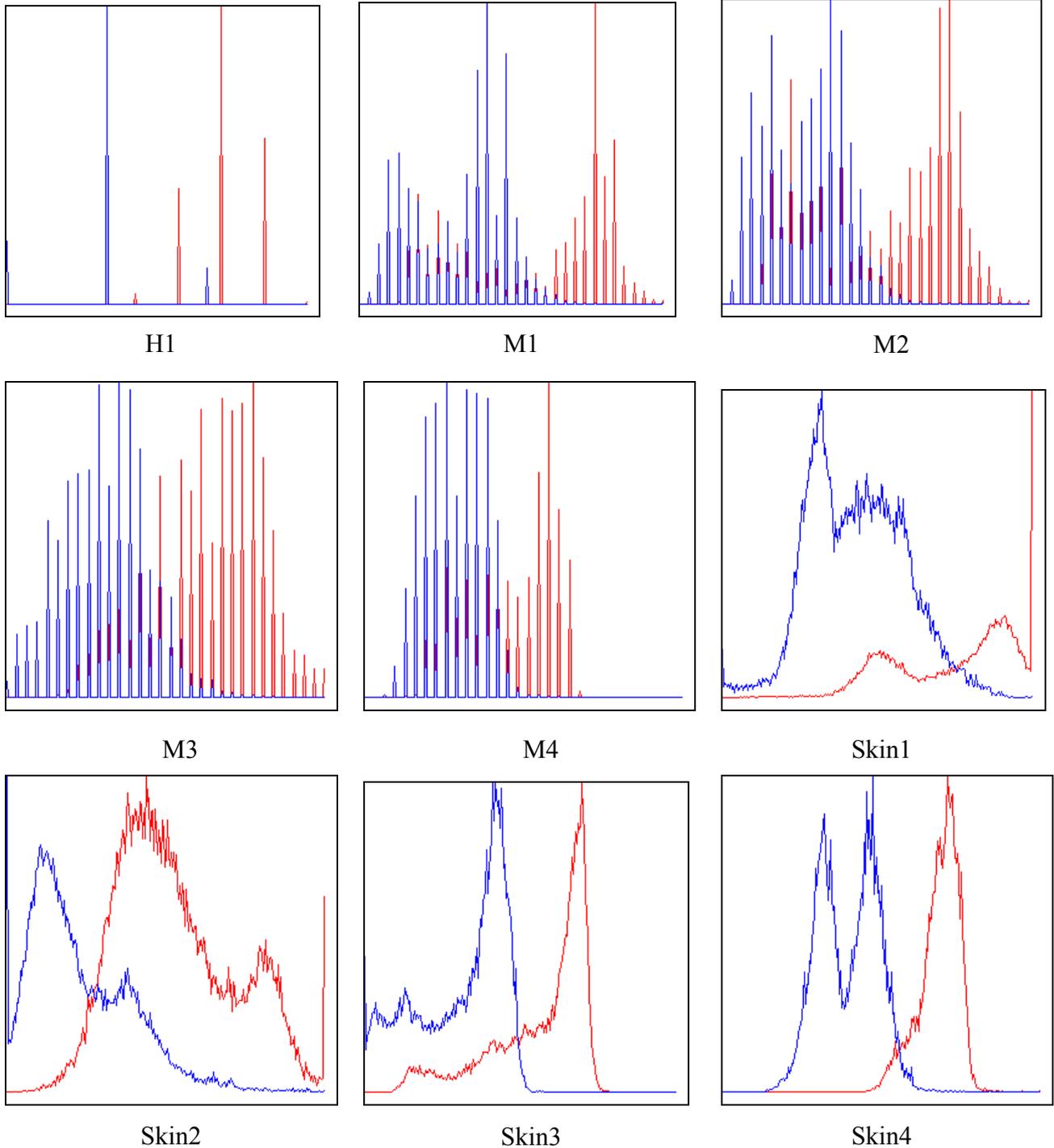
The choice of threshold values is, after many empirical tests, the best suited values.

The problem with these threshold values was to find the most accurate “common denominator” between all the pictures.

To back up (or sometimes to weigh) these results, the use of the red and blue histogram seems to me very interesting, as it is given only the amount of pixels of each colour, without any spatial distribution.

A question about the sample pictures is that Skin4 is a .bmp file, so it means that it should not have suffer from any compression, but once you have zoomed, you will see the typical squares from the Discrete Cosine Transform, processed by a JPEG encoding, so if a picture is encoded and reencoded in many formats, there will be changes in the picture at the microscopically level (even if we don't see too many changes with human eyes).

d. The histogram



As you can see, M1 and M2 have a similar histogram profile, the only difference is the offset. H1 doesn't have so much red and blue pixels, whereas Skin4 includes quite a lot of blue pixels and many red pixels.

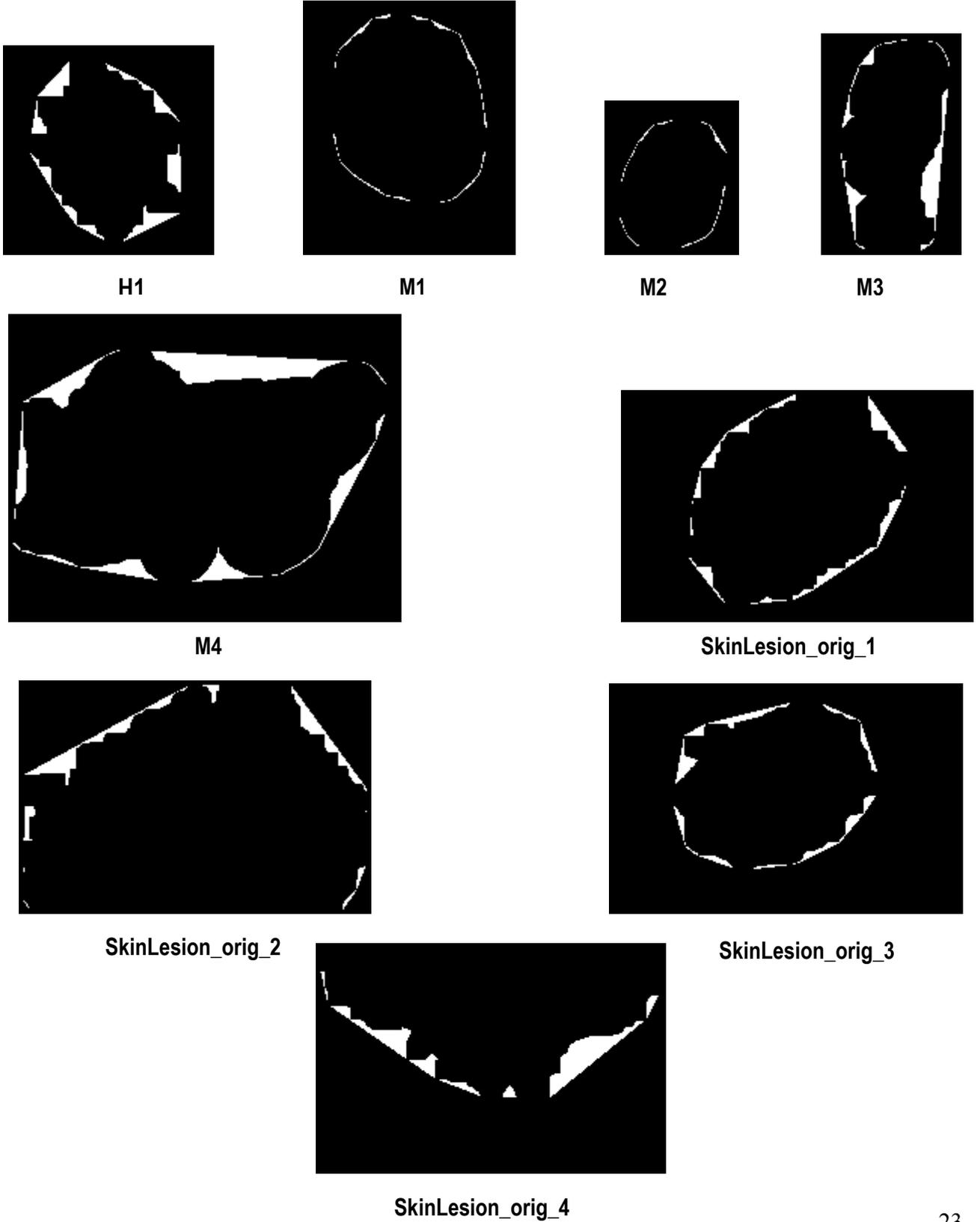
These histograms are directly processed from the original sample image, so it means that the skin around and outside the lesion is taken into account and this is not the case, as said before, for the detection of blue and red areas.

The amount of pixels considered as blue or red is processed thanks to the white pixel counter connected to the output image of each colours.

The results are :

	H1	M1	M2	M3	M4	Skin1	Skin2	Skin3	Skin4
Red	7461	821	224	3997	472	7665	6938	1507	14838
Blue	7611	1419	442	2639	5923	10659	4118	182	13576
Area	8077	10344	4871	7564	28211	14433	27713	11113	15592

e. The protrusions and indentations :



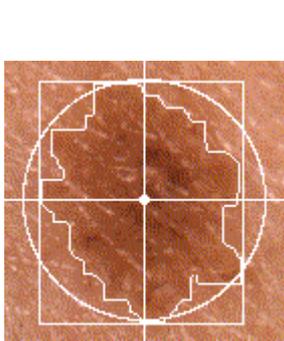
Final report for skin lesion project

Protrusions and indentations are suitable features in helping to the detection of skin lesions, even if they are not expressly mentioned in all the paper I read.

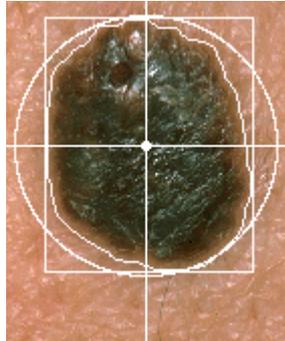
They can be linked to the compactness and circularity indexes.

M4 and Skin2 have quite a lot of protrusions, which are clearly highlighted.

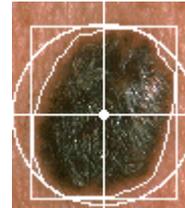
f. pictures with perpendicular axis :



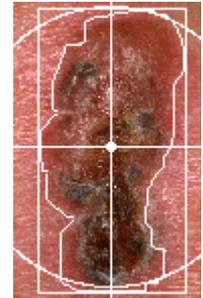
H1



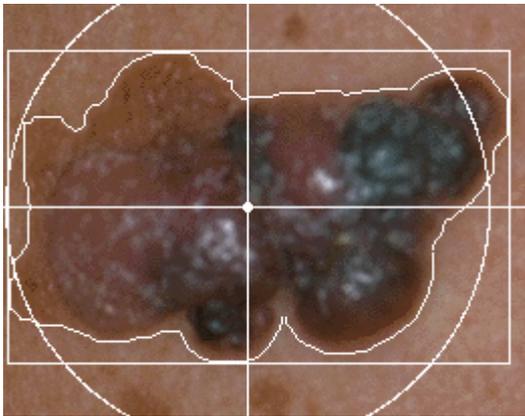
M1



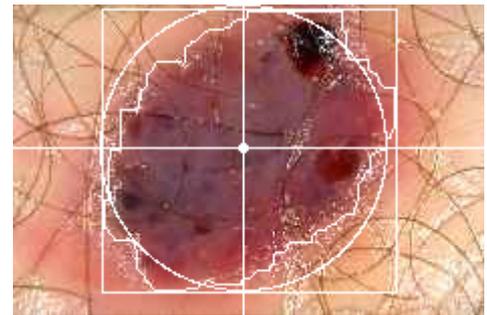
M2



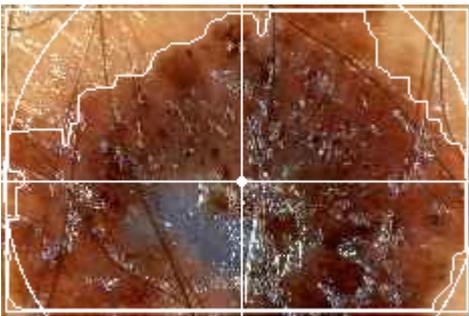
M3



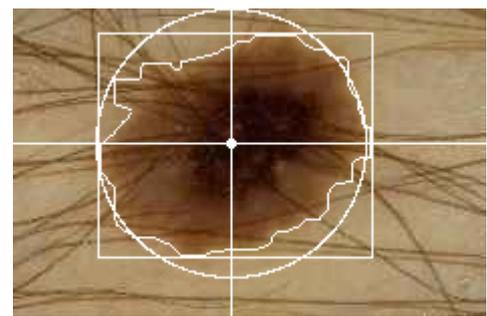
M4



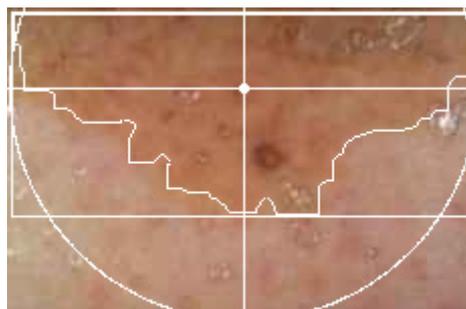
SkinLesion\_orig\_1



SkinLesion\_orig\_2



SkinLesion\_orig\_3

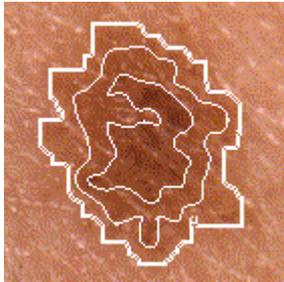


SkinLesion\_orig\_4

Final report for skin lesion project

These pictures are helping, like the first set, to target the shape of the lesion, by splitting the blob in 4 quarters to get a better idea of the appearance of the lesion and of the notion of distance. We can eventually see if a blob is centred or not in the lesion (Skin1, Skin2).

**g. The different shades of colours in the lesion :**



H1



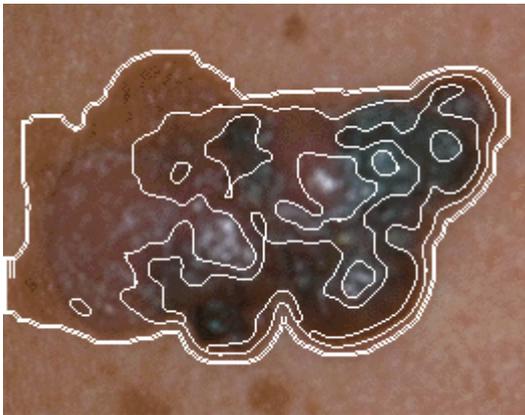
M1



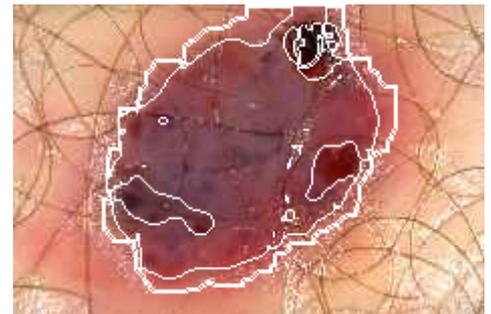
M2



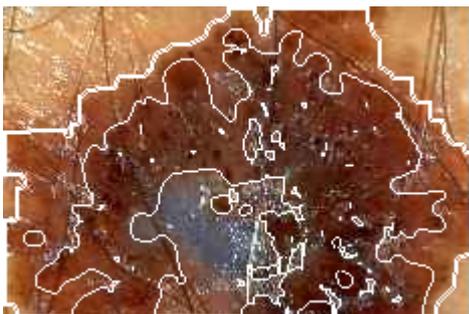
M3



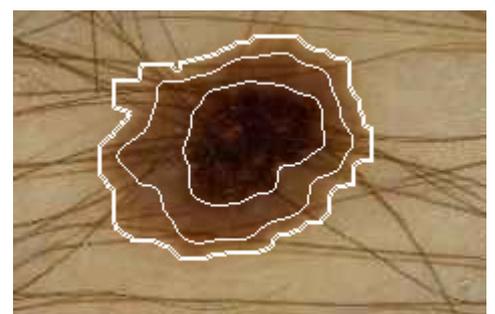
M4



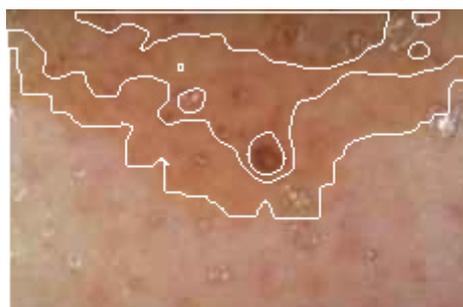
SkinLesion\_orig\_1



SkinLesion\_orig\_2



SkinLesion\_orig\_3



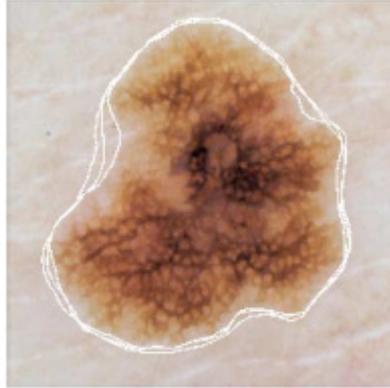
SkinLesion\_orig\_3

### Final report for skin lesion project

This set of picture is supplying a relevant view of the 3 main shades of colours and this is obtained thanks to the adaptive smoothing, a k means cluster and some post-processing blocks.

This set could be useful for the user, especially when he wants to draw manually the outline of the main visible shades but he can't because hairs are corrupting his track or when 2 juxtaposed shades are impossible to discern.

It is really a priceless tool, as dermatologists, when we ask them to draw the outline of the main colours for a few samples, have different drawings, even with the same dermatologist and the same image !  
cf [3] :



3 boundaries drawn by physician.

So this tool is helpful because it is drawing automatically and without hesitation the different outlines of the lesion, there is no human failure like stress, tiredness or inattention in the final result.

Glint is, even after the adaptive smoothing, still a bit affecting the detection on a few pictures (M2 and M4).

For M1, there is a tiny, uniform dark brown blob at 11, which is detected and causing the curled-up banana.

The crater of M3 is better defined with this view, as well as the other similar areas.

Darker blobs are quite well detected, for example for Skin1 and Skin4.

There is a beginning of over-segmentation for M4, but this is mainly due to glint.  
For Skin3, hairs have almost no more incidence on the colour-segmentation.

## IX/ Chapter 4 : Problems raised

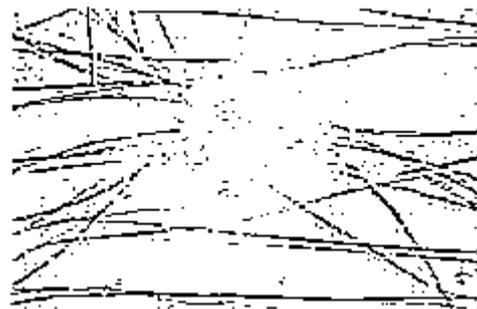
For the segmentation, my first approach was to try to isolate and suppress the hairs so that I would have subtracted them from this original picture and then applied a kind of interpolation to get back something close from the original image, but without hairs. The idea was that, if hairs are already removed from the input image, the result will obligatory be better for the rest of the process.

Actually I did some research about some possibilities to isolate and remove hairs, I found, among other things, a software called Dullrazor, from the British Columbia Cancer Agency. This program is really good for removing the hairs, even if there are a lot of them, but we can still detect where hairs were, cf [3].

A solution I first found interesting was an histogram stretcher (to emphasize the hairs) followed by switch to HIS domain. After that, we work on the Intensity component and we apply a adaptive threshold  $5 \times 5$  to finally get only the hairs, as shown on the following pictures for Skin1 and Skin3 :



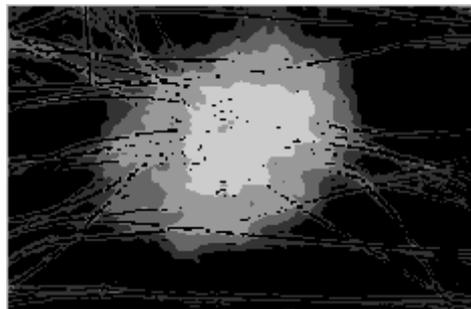
Skin1



Skin3

Here hairs are quite well isolated around the lesion but over the lesion there is almost no hair detected. The idea is then to subtract them from the original picture.

The next picture is an ADDing between a 4 grey-scale image and the bend of the hairs for Skin3 :



As you can see, there are still gaps between the detected hairs and the clustered image and once hairs subtracted, gaps are even bigger. These gaps were hard to remove and even to minimize with specific convolutions.

So I didn't go further in this direction, because, from the beginning of the project I didn't want to deeply alter the part of the image (that is containing the skin lesion) for the rest of the process.

For this project and more especially for the adjusting of the Adaptive Smoothing and for the reddish and bluish area detection, the trick was to find the most suitable values for these blocks, so that we obtain the most acceptable result. It is like finding the common denominator for a fraction. But in this case there is a part of trial and error.

I want to stress that I didn't care about noise in the image (salt&peper, Gaussian,...) because it would be "a disgrace" to have noise at this stage of the skin cancer detection ; I mean that noise can mainly comes from the acquisition process and absolutely not from the connection from the camera device to the computer. It is obvious that everything is done against noise prior to the data-analyse process.

As time performance could be important in the goal of a mass analysis of pictures databases, I tried as far as I could, to use blocks greedy of CPU (for example Reconstruction by dilation).

## **XI/ Chapter 5 : Possible and future improvements**

What could be great to sharpen this skin-lesion analyser, could be to work on the detection and highlighting of similar textures in the space, for example with textels that are forming most of the lesion. Image could be analysed to find similar colour or same structural element lattices.

An other improvement would be the dimming of glint, because this is sometimes a real feature we have to deal with. But if the input images are standardized (same resolution, lightning, angle and distance from the lens), this problem could be minimized.

## **XII/ Conclusion**

This solution is a non-destructive system, which brings out key features in order to help dermatologists in their diagnostics. This system is not a dermatologist, it is just a useful tool for people who know how to use it and how to interpret its results and this system is not giving a final answer to the question "is this lesion the beginning of a skin cancer".

The ABCD rules are quite well represented with the 17 different highlighted features. The results are quite correct as regard of what we can see in the literature [3],[4], but to my opinion, the final solution would be more sharpen and accurate if inputs images were standardized (format, resolution, light, colour panel) so that we will have known environment for each picture so each one can be treated the same manner to have a fair process, this has even more sense if this system is designed to help the dermatologist in his study or surgery, because it is unlikely that he will change the lens or the camera 10 times a day. In this viewpoint all the acquisition system will be defined once and for all.

It was really interesting to achieve this project, especially to apply the Machine Vision background to medical world.

## XII/ References

- [1] : “*Automated Melanoma Recognition*”, by H. Ganster, A. Pinz, R. Röhner, E. Wildling, M. Binder and H. Kittler.
- [2] : “*Initial results of Automated Melanoma Recognition*” by H. Ganster, M. Gelautz and A. Pinz, from the Institute for Computer Graphics, Technical University Graz, Austria.
- [3] : “*Towards a computer-aided diagnosis system for pigmented skin lesions*” , by P. Schmid, Joël Guillod, Jean-Philippe Thiran from the Swiss Institute of Technology.
- [4] : “*Colour segmentation for the analysis of pigmented skin lesions*”, by P. Schmid and S. Fisher from the Swiss Federal Institute of Technology.
- [5] : “*Segmentation of Skin Cancer Images – a report*”, by Padmapriya Nammalvar from the Dublin City University.
- [6]: *Dullrazor and the literature related to*, copyright to British Columbia Cancer Agency.
- [7] : “*A New Algorithm for Border Description of Polarized Light Surface Microscopic Images of Pigmented Skin Lesions*”, by C. Grana, G. Pellacani, R. Cucchiara and S. Seidenari.
- [8] : “*A Noninvasive Skin Imaging System*”, by Symon Cotton, School of Computer Science, University of Birmingham, UK.
- [9] : “*A Survey of Shape Analysis Techniques*”, by Sven Loncaric, from the Department of Electronic Systems and Information Processing of the University of Zagreb.

# Global view of the NeatVision file

