

Automatic nevi segmentation using adaptive mean shift filters and feature analysis

Michael A. King^a, Tim K. Lee^{ab*}, M. Stella Atkins^a, David I. McLean^c

^aComputing Science, Simon Fraser University, 8888 University Drive, Burnaby, Canada

^bBC Cancer Agency, 601 West 10th Avenue, Vancouver, Canada

^cVancouver Hospital and Health Sciences Centre, the University of British Columbia, Vancouver, BC, Canada

ABSTRACT

A novel automatic method of segmenting nevi is explained and analyzed in this paper. The first step in nevi segmentation is to iteratively apply an adaptive mean shift filter to form clusters in the image and to remove noise. The goal of this step is to remove differences in skin intensity and hairs from the image, while still preserving the shape of nevi present on the skin. Each iteration of the mean shift filter changes pixel values to be a weighted average of pixels in its neighborhood. Some new extensions to the mean shift filter are proposed to allow for better segmentation of nevi from the skin. The kernel, that describes how the pixels in its neighborhood will be averaged, is adaptive; the shape of the kernel is a function of the local histogram. After initial clustering, a simple merging of clusters is done. Finally, clusters that are local minima are found and analyzed to determine which clusters are nevi. When this algorithm was compared to an assessment by an expert dermatologist, it showed a sensitivity rate and diagnostic accuracy of over 95% on the test set, for nevi larger than 1.5mm.

1. INTRODUCTION

Nevi density is the strongest predictor of malignant melanoma, with about 50% of malignant melanoma originating from pre-existing nevi [1-3]. Therefore, nevus counting is often an integral part of many studies that look at malignant melanoma [4-7]. Since manual counting is often costly and inaccurate, an automatic method of segmenting and counting nevi would help standardize studies. An automatic method would also aid in the counting and registration of patients who are at a high risk of developing malignant melanoma and need to get their nevi examined frequently.

Counting nevi accurately and consistently is a difficult problem, even for trained dermatologists. Nevi are easily missed if they are small, and other skin conditions such as freckles can be mistaken as nevi. Doing automatic nevi segmentation is no less difficult, with varying light conditions one cannot rely on color to correctly segment the image. Different colored lights change the color of the nevi in a non-linear way so that the original color under natural light cannot be recovered. Another problem is that nevi can be very large, or be in a non-elliptical shape, such as when two nevi close to each other merge; therefore shape and size are not reliable indicators as well. Since there is no one measure that can reliably classify nevi, a feature space approach is taken. Each relevant feature, such as size, is taken into account, and a feature vector is constructed.

The algorithm described in this paper consists of three major steps; the first step is an iterative algorithm that uses an adaptive mean shift algorithm that removes noise from the image while preserving the nevi. The second step is a simple merging algorithm that creates large clusters of pixels and identifies the possible nevi in the image. The final step classifies each candidate as nevi or not using feature vectors and a feature space.

The mean shift algorithm used, unlike the original algorithm [8], [9], [10], is not guaranteed or necessary to converge. Each pixel is replaced with a weighted averaged of pixels with similar intensity to itself in a small neighborhood. This averaging is done iteratively, typically 4-6 passes will smooth the image sufficiently, and any additional iteration will result in small differences only.

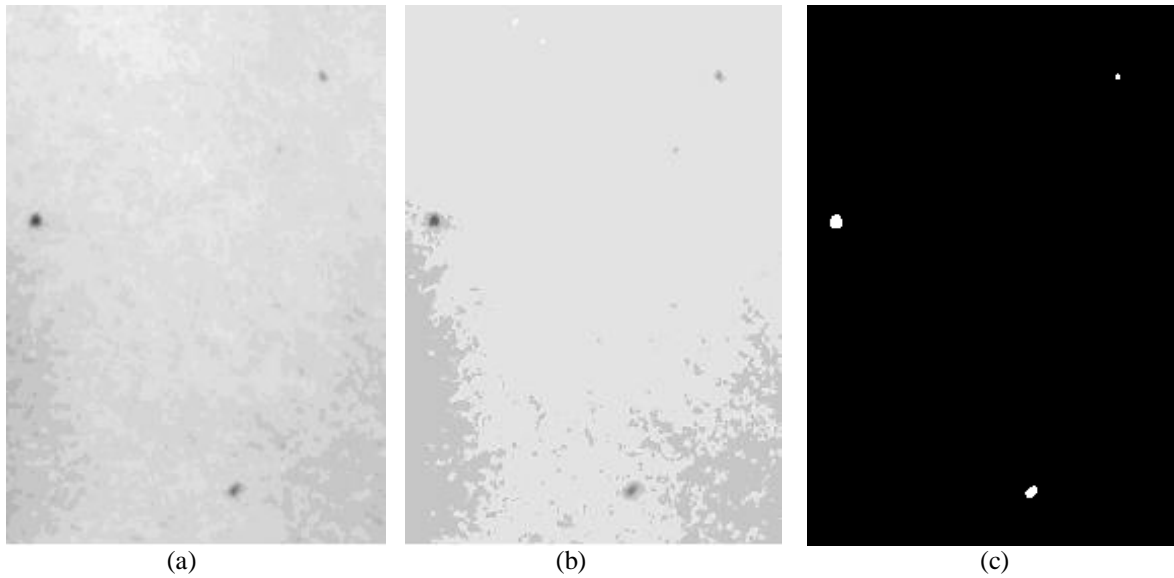


Figure 1. a) original skin image b) after mean shift c) candidates selected

A consistent feature of nevi is that they always have lower intensity than the surrounding area, so nevi appear as dark spots in the image. Once the image is smoothed from the mean shift algorithm, any local area minimum in intensity is marked as possible nevi.

Figure 1a) shows section of a skin patch used, figure 1b) shows the same portion of the skin that has been filtered by the mean shift algorithm. Figure 1c) has the areas that are local minima, but not touching the boundary of the image, highlighted.

2. MEAN SHIFT

In this section the general mean shift algorithm will be discussed [8], as well as common extensions [8][9][10] and our specific extensions used for the expressed application of nevi segmentation. Mean shift has been used for image segmentation as well as video tracking [10]. It is known as a good clustering algorithm that preserves edge boundaries as well keeping a representative color or intensity of each cluster.

Let S be the set of pixels in the image to be processed. Let x be the element in S that is currently being processed and $I(x)$ be the pixel intensity of x . Let K be the kernel of the mean shift filter. A flat kernel is defined as:

$$K(z) = \begin{cases} 1 & \text{if } z \leq a \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where a is the radius of the neighborhood in intensity space. The sample mean is defined as:

$$m(x) = \frac{\sum_{s \in S} K(\text{abs}(I(s) - I(x)))w(s)I(s)}{\sum_{s \in S} K(\text{abs}(I(s) - I(x)))w(s)} \quad (2)$$

where $w(s)$ is a weight function. In the mean shift filter, each pixel value is replaced with its sample mean value, with all pixels in the image updated simultaneously, and the filter is applied several times; so each iteration $I(x)=m(x)$ for all $x \in S$. The mean shift algorithm is a gradient ascent algorithm, which converges into stable clusters when using the flat kernel.

2.1. Adaptive Kernel

Typically, the mean shift kernel in any application is consistent and does not change, as well as having a kernel that guarantees convergence. Two common kernels are the flat kernel, from equation (1) as well as the Gaussian kernel:

$$K(z) = \left\{ e^{-z^2} \right\} \quad (3)$$

For practical uses, the Gaussian kernel is not used, since each pixel has the whole image as a neighborhood. A mixture of a flat kernel and Gaussian kernel can be used, called a truncated Gaussian kernel:

$$K(z) = \left\{ \begin{array}{ll} e^{-bz^2} & \text{if } z \leq a \\ 0 & \text{otherwise} \end{array} \right\} \quad (4)$$

A problem with this kernel is that it treats everything the same, in that it will preserve all objects within the image if there are enough pixels in its neighborhood with similar intensity. If there are some patches of hair in the image, it would leave the hair in the image, while averaging the intensity of the hair. We do not want this behavior as it could produce false positives in the results. By modifying the kernel to adapt to different local histograms, we get:

$$K(z) = \left\{ \begin{array}{ll} e^{-B(Y)z^2} & \text{if } z \leq a \\ 0 & \text{otherwise} \end{array} \right\} \quad (5a)$$

where Y is the set of all the pixels in the neighborhood of x , so $Y \subseteq S$, where the function $B(Y)$ is defined as:

$$B(Y) = c_1 \left(\sum_{y \in Y} f(\text{abs}(I(y) - I(x))) / |Y| \right) \quad (5b)$$

where c_1 is a constant and $|Y|$ is the number of elements in the set.

$$f(z) = \left\{ \begin{array}{ll} 1 & \text{if } z \leq a \\ 0 & \text{otherwise} \end{array} \right\} \quad (5c)$$

As $B(Y)$ decreases, it resembles a flat kernel. Flat kernels tend to smooth out small objects, a few pixels that are outliers but are still within a will push the pixels toward the intensity of the majority of the pixels in the neighborhood. A few pixels of a hair in an image will have a low $B(Y)$ value, since typically there is mostly skin in the background. At least a few pixels of the skin near the hair will be in its neighborhood and intensity range, so these pixels will change the intensity toward the skin intensity on every iteration, merging the hair into the skin.

Nevi usually are composed of many pixels in a small area, so $B(Y)$ will be higher since there are many pixels with similar intensity close together. A high $B(Y)$ will result in a high standard deviation Gaussian kernel which doesn't allow the skin pixels to influence the intensity of the nevi pixels as much, since pixels farther away from the pixel intensity will be weighted much lower. Figure 2 shows how a flat kernel tends to merge objects into the background, while a Gaussian kernel tends to preserve the objects better.

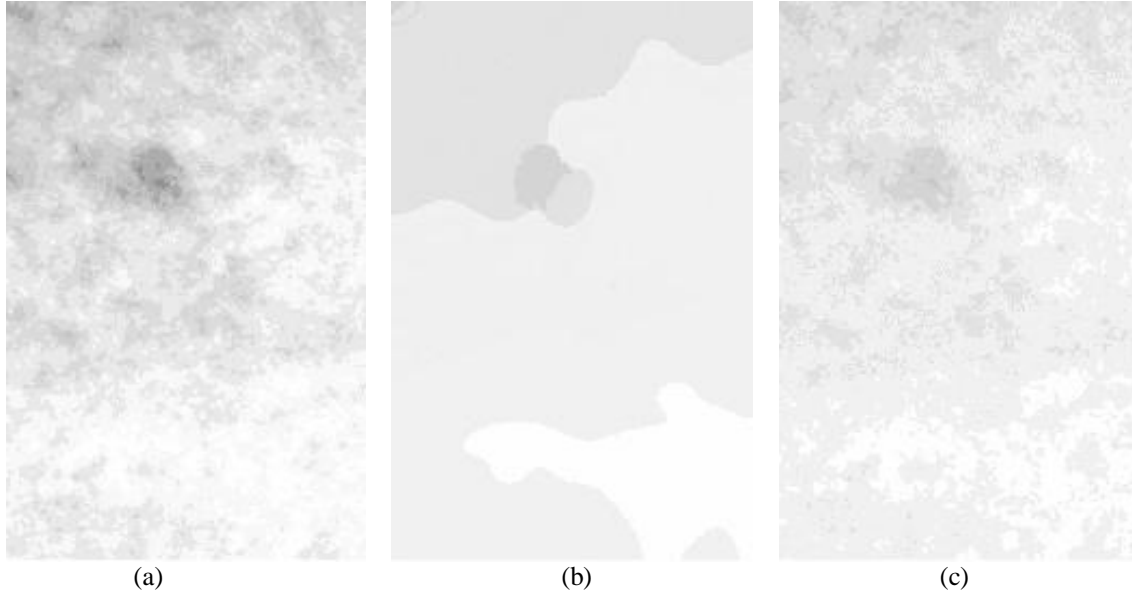


Figure 2. a) original skin image b) mean shift – flat kernel c) mean shift – Gaussian kernel

2.2. WEIGHTING FUNCTION

The weighting function, $w(s)$ of equation (2) is usually fixed throughout all iterations or set for each iteration. Typically it is used to increase the speed of convergence, with a higher $w(s)$ the pixel typically moves more in intensity space with each iteration. It could also be set to decrease with every iteration, to make sure it doesn't jump over the true mean. In this situation, we are more interested with eliminating hairs and small noise and not speed of convergence. Let the weight be defined as:

$$w(s, x) = \frac{c_2}{D(s, x)} + c_3 \quad (6)$$

Where c_2 and c_3 are constants, and $D(x, s)$ be the spatial distance between pixels x and s . Since nevi are always solid in shape and without holes, giving pixels nearby higher weight will help the edges of the nevi stay within the nevi cluster.

2.3. ADAPTIVE MEAN SHIFT ALGORITHM

After we include the weighting function, the modified mean shift algorithm is:

$$m(x) = \frac{\sum_{s \in S} K(\text{abs}(I(s) - I(x)))w(s, x)I(s)}{\sum_{s \in S} K(\text{abs}(I(s) - I(x)))w(s, x)} \quad (7)$$

The sample mean, $m(x)$, is calculated for each pixel for the whole image. Then the pixel intensity value for each pixel is updated with its sample mean value, $I(x) = m(x)$. It was determined experimentally that after 5 to 6 iterations the results of the mean shift did not change the image significantly anymore.

2.4. COMPLEXITY

The mean shift algorithm's speed is dependent on three factors: number of iterations, size of neighborhood and size of the image being processed. The number of iterations can be set to a small constant, such as six, but in standard practice is set to iterate until convergence. The size of the neighborhood is difficult to choose, as a small number will create many small clusters and be subject to noise, since each cluster has no chance to merge with pixels farther away from the pixel. If a large neighborhood is chosen the quality of the results increase, but the run time increases in polynomial time with the radius of the neighborhood. As the size of the image increases, better results are found, since the nevi will occupy more pixels in the image, as opposed to a smaller, lower resolution image.

The number of calculations for the mean shift is therefore:

$$\text{Number of Calculations} = \text{Iterations} * (\text{area image}) * (\text{area of neighborhood}) \quad (8)$$

We developed a variation of the mean shift algorithm to produce similar results, while reducing the number of iterations to calculate the sample mean to less than two.

Instead of moving the pixel in intensity space, we want to move it spatially first. We use a more basic form of the sample mean equation:

$$m(x) = \frac{\sum_{s \in S} K(\text{abs}(I(s) - I(x))D(x, s))}{\sum_{s \in S} K(\text{abs}(I(s) - I(x)))} \quad (9)$$

which is a modified form of equation (2) with the weighting function set to one. Also the kernel used is the flat kernel, defined in equation (1). The main difference is that instead of doing the calculations in intensity space, the calculations are done spatially. So the values of s and x represent the position of the pixels. This must be done in two dimensions, the X and Y axis. The algorithm is as follows:

1. For each pixel, calculate the spatial sample mean
2. For all pixels with sample mean zero have their intensity sample mean calculated as in equation (7)
3. For each pixel, while the spatial sample mean is not zero, current position = current position + sample mean
4. Set the pixel to the intensity sample mean of the pixel that step 3 ends on

The center of clusters should have a spatial sample mean of zero, so all pixels in the object take on the value of the intensity sample mean of the center pixel in the cluster.

3. CANDIDATE SELECTION

After the image has been processed using the mean shift algorithm, an image with many clusters of pixels with the same value will be produced. Pixel values are then quantized, to form larger clusters. Pixel values can be quantized into 64 or 128 values, from the initial 256, while producing good results. The clusters having a lower intensity than any of the other clusters surrounding it, and that are not against the border of the image, are identified as possible nevi candidates.

The second part of this step is to find the total area that the nevus occupies. Often in a large nevus, the center is significantly darker than the outer areas of the nevi; this results in the cluster only representing part of the nevus. An iterative procedure is performed to attempt to merge outer nevi clusters to the central cluster, to compensate for underestimating large nevi..

4. FEATURE SPACE

Once the possible nevi are found, statistics for each is generated resulting in an associated feature vector. Based on labeled data by an expert dermatologist, the feature space is divided into nevus and not nevus. Any candidate found within the feature space indicating it is a nevi is indicated as such, otherwise it is discarded. For each candidate found, several features are calculated. The most important features are size and contrast. Another supporting feature that is used is shape, in terms of how elliptical and elongated it is. Size is defined as the area; the number of pixels in the candidate. The contrast is defined as the average intensity of the pixels in the candidate divided by the average intensity of the pixels neighboring the border of the candidate. The elliptical measure is defined as the number of pixels within its convex hull divided by its size. The elongation is the ratio between minor axis and the major axis of the nevus.

The minimum size depends on the resolution of the image as well as the distance of the subject to the camera. A minimum size of 1.5mm^2 is enforced, as smaller candidates could simply be noise. Very large nevi are also unlikely, but cannot be ruled out. The contrast is a very important feature; the darker it is, the more likely it is a nevus, and not a dark skin patch. The elliptical measure has a desirable value of one, which could mean a perfect ellipse. But two nevi that have been merged would also have this measure equal to or close to one. As this measure decreases, it is more unlikely to be a nevus. The elongation of a perfect circle would be equal to one. Setting a limit on the elongation and ellipticity prevents errors from shadows, which are typically long, thin and dark which are generated by the shoulder blades. Our program hypothesizes that a nevus is 2-5mm wide, dark, and a perfect circle. As candidates deviate more from this ideal, they are less likely to be nevi.

Each candidate has an associated feature vector, based on the above features. Feature vectors from expertly labeled nevi on images forming a training set from this class of images were found and were analyzed. The feature space was then divided into two regions, corresponding to nevi or not nevi, with the nevi region enclosed in a four dimensional sphere in the feature space. When a new candidate is introduced, it is classified based on what region it falls into.

5. RESULTS

Application of the algorithm on a test set of eight images taken from a 1993 nevus counting study showed promising results. The test set consisted of large sections of skin from the backs of children taken at a resolution of approximately 16 pixels/ mm^2 . The test set was taken with a standard 35mm camera and digitized for input into the algorithm. All images were taken in a similar setting; however, some images were taken where the skin was not perpendicular to the lens of the camera, causing shadows from the body and shoulder blades. Furthermore, the skin images were taken from a range of different ethnic groups with varying skin color.

All eight test images were labeled by an expert dermatologist. When compared with the nevi over 2.25mm^2 in size that were chosen by the dermatologist, the algorithm found 54 true positives, 1 false positive and 3 false negatives which results in a sensitivity rate of 95% and a diagnostic accuracy of 98%. No nevi above 2mm in diameter in size were missed by the algorithm. The neighborhood used in the tests was eight pixels in both x and y directions, corresponding to 4mm^2 , with six iterations of the mean shift algorithm. Experimentally, the most reliable measure was the contrast between the intensity of the nevi and the outer skin area, with the average nevi intensity approximately 20% darker than the surrounding skin area.

Common false positives were from hair covered areas. These are not filtered out using the mean shift filter, since they are clustered together, so if they happen to have a similar shape as a nevus, hairs would be classified as nevi. Nevi lying near the border of a shadow or near body folds were particular difficult to find as well, since they contrasted poorly with the skin area, or the intensity moved toward the color of the shadow.

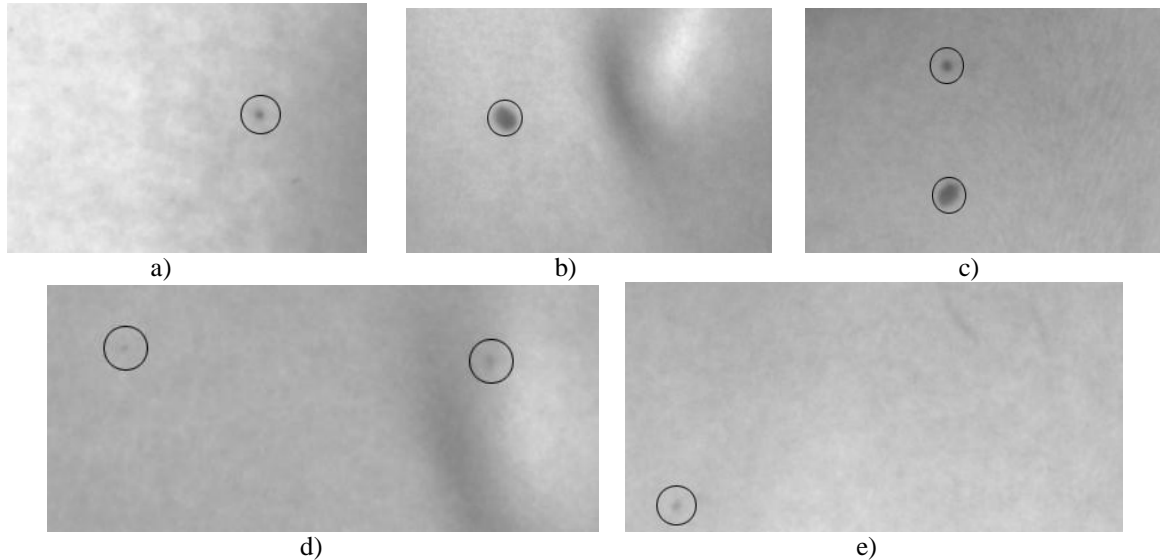


Figure 3. Sections of skin with found nevi circled

6. CONCLUSION

A method to automatically segment nevi from skin images is presented in this paper, using an adaptive mean shift filter. This filter has the advantage that no prior knowledge is needed about the number of nevi clusters present on the skin because decisions are made locally in a small area. Another advantage of the algorithm is that it is robust in varying lighting conditions, so it can potentially be used in many different situations. Results so far are very promising, as the mean shift filter merges small objects such as hair into the skin, making nevi easily stand out from the skin. The mean shift algorithm, unlike many other de-noising algorithms, preserves the shape and intensity of the nevi, allowing for accurate statistics to be gathered. The identification of nevi also correlates well with the eight images manually labeled by an expert dermatologist.

7. REFERENCES

1. C. D. Holman and B. K. Armstrong, "Pigmentary traits, ethnic origin, benign nevi, and family history as risk factors for cutaneous malignant melanoma," *Journal of the National Cancer Institute*, vol. 72, pp. 257-266, 1984.
2. E. A. Holly, J. W. Kelly, S. N. Shpall, and S. H. Chiu, "Number of melanocytic nevi as a major risk factor for malignant melanoma," *J. Am. Acad. Dermatol.*, vol. 17, pp. 459-468, 1987.
3. T. M. Skender-Kalnenas, D. R. English, and P. J. Heenan, "Benign melanocytic lesions: risk markers or precursors of cutaneous melanoma?," *Journal of the American Academy of Dermatology*, vol. 33, pp. 1000-1007, 1995.
4. D. R. English, R. MacLennan, J. K. Rivers, J. Kelly, and B. K. Armstrong, "Epidemiological studies of melanocytic nevi: protocol for identifying and recording nevi," International Agency for Research on Cancer, Lyon, France IARC Internal Report 90/002, 1990.
5. R. P. Gallagher, J. K. Rivers, T. K. Lee, C. D. Bajdik, D. I. McLean, and A. J. Coldman, "Broad-spectrum sunscreen use and the development of new nevi in white children: a randomized controlled trial," *The Journal of the American Medical Association*, vol. 283, pp. 2955-2960, 2000.
6. T. S. Wiecker, H. Luther, P. Buettner, J. Bauer, and C. Garbe, "Moderate sun exposure and nevus counts in parents are associated with development of melanocytic nevi in childhood: a risk factor study in 1,812 kindergarten children," *Cancer*, vol. 97, pp. 628-38., 2003.
7. T. K. Lee, M. King, S. Lau, M. S. Atkins, and D. I. McLean, "Automatic nevus counting: A pilot study," *Skin Research and Technology*, vol. 9, pp. 191, 2003.

8. Y. Cheng, "Mean shift, mode seeking and clustering", IEEE Trans. PAMI, Vol.20(6), pp. 790-799, 1995.
 9. Dorin Comaniciu and Peter Meer, "Mean Shift: A Robust Approach Toward Feature Space Analysis, IEEE Trans. PAMI, Vol.24(6), pp. 603-618. Date?"
 10. D. Comaniciu and P. Meer, "Mean shift analysis and applications", Proc. 7th Intern. Conf. Computer Vision, Greece, pp. 1197-1203, 1999.
- * tlee@bccancer.bc.ca; phone (604) 707-5922