Automated Analysis and Discrimination of Carcinoma in situ of Thermographic (IR) Images

Brad Morantz,
Imagery & Technology Systems Division
Science Applications International
Tucson AZ USA

Abstract

Cancer has become a major health concern world wide. Early detection has been shown to save lives. Imaging infrared, originally a military technology, was applied to medical imaging several decades ago. It is based on the concept that the additional activity in the cells translates to extra heat, which in turn causes more infrared photons to be emitted.

Several years ago when this concept was first applied to medical imaging, computers were much slower and memory was limited and expensive. With the advent of more powerful computers and less expensive and more plentiful memory, as well as technological advances in detectors, the study of infrared medical imaging has renewed.

The first step is to determine if the person has a carcinoma, and the second step would be then to locate and identify it. Some thermographic images were statistically analyzed in the attempt to discriminate between images with a tumor and those without. The images were first processed by standard digital image processing techniques, and then statistical analysis was performed. On the limited sample images that were available, this method was successful at classification of malignant versus benign human issue.

Introduction

Early detection of breast carcinoma has been a subject of much research. It is expected that 12.5% of females will experience this some time in their life, and that 20% of these will die from it. [Ng et al, 2002] [ACS, 2000] Breast cancer in females is a leading cause of death in the USA. It is believed that earlier detection of the carcinoma can improve the prognosis of survival [Kieth et al, 2001]

Currently, the accepted method for detection is radiographic mammography. Using Xrays, the breast is physically manipulated (quite painfully to the patient) and images are created in several angles. Digital x rays, where a semiconductor receptor replaces the film, are more sensitive than film and allow for a smaller exposure to radiation. Unfortunately, these are not widely implemented, due to the high cost.

A new advance on this procedure is 3D tomographic x rays, where 11 images are taken at various preset angles, and a full 3D image is calculated and produced by a computer. [**reference**] This process is very new and is yet to be accepted by the FDA. Additionally, it does cause extensive radiation exposure.

These radiographic methods are sometimes blamed for the increasing incidence of this malady. [Darby & de Gonzalez, 2004] [Gotzsche & Olsen, 2000] In *Politics of Cancer,* Samuel Epstein discusses the cumulative carcinogenic effects of radiation on premenopausal breast tissue. [Epstein, **YEAR**] The physical manipulation of squeezing the breast between the two plastic plates is not only painful for the patient, it has been reported to cause the cancer cells to metastasize. [Goldberg, **YEAR**]

Other standard methods are magnetic resonance imaging (MRI), ultrasound, and positron emission tomography (PET). Thermography is an older process and has had a renaissance. As of this writing, health insurance companies do not endorse this method, so it has not gained wide acceptance. Near infrared transillumination scanning is an old concept but the resolution was poor. New technology has allowed increased resolution, but this method still requires the physical manipulation of squeezing between the glass plates. [reference] An even newer and yet fringe method has been experimented where a dog is trained on the breath of people with and without invasive breast malignancies. After a training period, in one test, the dog correctly identified 98% of the test cases.

An advantage of the last two methods is that they are totally passive and without any possible side effects. There is some research that shows that certain women carrying the A-T gene are even more susceptible to a malignancy upon exposure to radiation. [Epstein **year**]

Another advantage of the dog method is that it is fast, in that a person can get a quick yes or no answer, and for those cases where it was the former, then further testing can be done to verify the results and locate the tumor. One would wonder if this method works for in situ tumors or just invasive ones where the cells have traveled throughout the body and are present at the aveoli and are therefore expressed into the breath where the dog can sense it. There are also reports that dogs can sense the onset of an epileptic seizure. But until these animal methods are more substantiated in made institutional, we must look at more traditional methods.

What we can learn from the dog example is the concept of having a totally passive, non painful, near instantaneous test that indicates the existence or not of a carcinoma. Thermography, as will be discussed in this paper, presents just such an opportunity. Further, the thermograph can be taken and first a quick test for the existence of neoangiogenisis can be done. If there is need for further analysis, then the image can then be processed by a more extensive analysis such as assymetrical (chinese guy) Thermography, 3D tomography, or another more extensive method. The advantage is that this is an initial test, with no potentially dangerous side effects, that can determine the need to do any more testing. As only 12.5% of women get this, then 87.5% do not, and this would create financial savings to the health care system, and not expose women to additional risks. This is an application of Ockham's razor.

Current screening methods are somewhat effective for moderately fast-growing tumors that may or may not be metastastic by the time of their detection. [Kieth et al, 2001] Radiography and ultrasound look for anatomical variation and structural distinction of the tumor from the surrounding tissue. [Ng et al, 2001] There has been some work in looking for assymetry, both with radiography and thermography. [Qi et al] Qi et al uses edge detection and hough transforms to partition the image and then look for dissimilarities in heat patterns. This method requires more extensive calculations and analysis. The method described here is what machinists call a “go or no-go test”, in other words, first to determine if there is something abnormal there, and if not, it is not necessary to continue analyzing.

These methods are not even used for those under 40 years of age, unless other symptoms (lesions, lumps, etc.) are noted. Some research has shown that all breast cancers actually begin in premenopausal women. This was based upon a study of 500 consecutive cases in 1988. [Simpson in Kieth] Radiography is not recommended in women of child bearing age because the definition is relatively poor due to the normally high tissue density of the breast parenchyma. Self palpation, if done correctly and regularly, fails to detect masses under 1 cm in diameter. A mass of this size is approximately 1 billion cells. Therefore, a lump found by self-palpation is often large and metastatic. [Kieth et al, 2001]

Thermography, or infrared heat scans have been used for imaging many things, from military applications, to home insulation contractors, to medical diagnosis. Thermography captures the natural heat radiation from any object that is above absolute zero (0 K or -273 C). The Stefan-Boltzmann Law defines the relation between radiated energy and temperature by stating that the total radiation emitted by an object is directly proportional to the object's area and emissivity and the fourth power of its absolute temperature. The emissivity of human skin is extremely high (within 1% of that of a black body). [Amalu, 2002] It is passive (no energy is transmitted to the object as in X-ray radiography) and it is noninvasive. It provides means of further investigating physiological situations on another dimension, without any possible adverse affects upon the patient. Neoangiogenesis, formation of new blood vessels to supply the growing malignant cells generates heat that can be detected with thermography.

New thermography applications are also being investigated in projects sponsored by the US Department of Defense using military thermal surveillance tools adapted for cancer detection. Both are enhancements of older thermal imaging technology based on the principle that heat equates to unwanted activity, in the case of breast cancer, abnormal cell proliferation. It also detects other activity such as a hematoma.

Thermography enables one to see the temperature pattern and look for abnormality. [Ng & Sudharsan, 2004] In a thermogram there is no radiation risk as it only captures the infrared radiation from the skin. It is totally painless and there are no adverse effects. The resulting image is a 2D map of the real 3D temperature distribution. [Jones & Plassmann, 2002]

Chemical and blood vessel activity is higher in precancerous tissue and in the area surrounding a developing carcinoma. This additional activity can cause a temperature difference between healthy and cancerous tissue. The rapidly growing malignant tumor needs an ever increasing supply of nutrients. Blood flow is increased to this area via a process called neoangiogenesis, the creating of new new blood vessels. [Ng et al, 2002]

The temperature and the vascularization of the breast, which are modified by endocrine, inflammatory, and tumoral influences, can be studied using thermograms, pictures created by temperature measurements made using infrared scanning. These thermographic examinations must be performed in a draft-free, temperature, and humidity controlled room, where a constant temperature 20 degrees C is maintained. The greatest interest of clinical thermography lies in its use in the detection of carcinoma. [Nyirjesy, 1982] In one study, tissues with a carcinoma were 0.85 C warmer on average than non cancerous tissue. [Ng et al, 2002] Commercially available uncooled microbolometers are available with 50 milliKelvin (mK) resolution, adequate to detect the variation required.

Contrast can be increased by applying a cool breeze so that the human autonomic system will respond by reducing blood flow to that area. The tumor is not controlled by the autonomic system and therefore will not have any reduction in flow and hence non attenuated heat generation. In a test of 85,000 patients, this method decreased the false-positive rate to 3.5% (96.5% sensitivity). [Gros & Gautherie, 1980]

One method of detecting this infrared radiation is in using arrays of quantum well infrared photodetectors (QWIP). Other detectors are made from mercury, cadmium, and tellurium or indium and antimony. All of these operate at 77 K. Earlier units accomplished this environment with liquid nitrogen cooling, but now use electric powered cryocoolers.

Subsequent analysis and interpretation by image analysis is the obvious next step. Due to processing limitations, thermography was not widely used until computers, memory, and image processing techniques had advanced adequately. One source reports that thermography was shelved in 1975 because of this limitation in ability to process and evaluate. [Ng et al, 2001] In this time while waiting for processing power to reach needed performance, the technology of infrared detectors and focal plane arrays (FPA) has advanced and the cost has attenuated, making the images more accurate and the hardware more readily available. As of this writing, resolution of 10 millikelvin (.010 Kelvin) has been achieved.

Artificial neural networks have been employed successfully for the classification of these thermographic studies. [Ng et al 2001, Ng et al 2002] Other technology such as support vector machines (SVM), naïve Bayes classification, and clustering techniques (statistical, hierarchical, and self organizing map) further provide dichotomous classification as required for this early screening process.

Methodology

As we do not have access to a high resolution thermographic scanner, we obtained images of both healthy and malignant subjects from wherever we could. Many labs and research facilities performing thermographic scans were contacted, but they all replied that they could not send us any images because of the new HIPAA regulations. Even torso portions with all identification removed were not available because of this new law and the current litigious nature in the USA.

Images were typically found in JPEG format and Matlab was used first to convert the discrete cosine transformations (DCT) into their respective 3 layer intensity matrices of red, green, and blue (RGB). The JPEG format has the intensity of each pixel defined by an 8 bit byte, allowing integer values from 0 to 255. This was divided by 255 and converted to double precision floating point number ranging from 0 to 1.0. Matlab requires this format to obtain a viewable image.

The initial step was to eliminate background so as not to include this unrelated data into our calculations. Removal consisted of setting the intensity to zero and calculating the total number of pixels in the foreground or image. In removing the value of these pixels, even though they are still in the electronic image, they no longer contribute to the statistical analysis. This was accomplished by counting the number of pixels that were classified as background and resetting the intensity of each to zero. At the end of the background classification process, this count was subtracted from the total number in the picture, resulting in the number of pixels in the image (foreground) only.

As the images were 8 bit uint format, the simplest method to predict correct pixel classification was exhaustive search over a narrow range. This did vary from image to image. Some images were very clean, and by examining pixel by pixel, the number that had intensities for all colours less than or equal to a threshold value were counted and considered to be background. All others were then foreground.

As the possible values are discrete numbers between 0 and 255, it became a simple search. A loop was created that varied the value of a threshold and then created a binary image, where all intensities below the threshold were set to zero and all others; one. The resulting images were examined. White spots on black or vice versa were potential errors. The foreground in binary mode should appear white, and block spots in it are indications of a pixel being misclassified. An optimal threshold level was set by choosing the value that had the smallest number of errors. This process is not universal so a better method was required.

In performing the binary conversion and observing the resulting images, it was noticed on one image that there were parts of the image that some pixels had a zero value for one colour, yet it was definitely in the foreground. A new classifier was programmed, again going pixel by pixel, but looking for pixels where intensities of each colour for that pixel were also below a threshold level and the sum total of intensities was below a different threshold.

Again, converting the image to binary and looking for white spots on the black background or black spots on the white image the simplest method for counting misclassifications. As the threshold level was varied, we expect to find an optimum value. When the misclassifications are plotted against threshold level, a concave upward parabola is expected, and the selected value should be at the very bottom.

The first statistic to be calculated is the percent of the foreground image that consists of each colour. This calculation is done by dividing the number of pixels in the foreground of each colour by the total number of pixels in the image. The three ratios obtained should add up to one (or very close).

The next step is to examine the intensity distribution for each colour. Again, the literature would lead one to expect that the image containing the carcinoma would have a heavier distribution in the higher intensities of red. This process is accomplished in Matlab by sorting the intensities of each pixel into statistical bins, each of the same size. It is easy to use 10 bins for each colour, providing a total of 30 bins per image.

Hypothesis

The independent variables are: image and health of patient. The dependent variables is: amount of heat emitted at each location/pixel. As the goal is to discriminate between images of those with cancer and those without, the null hypothesis is that all images have the same distribution of colour intensities within a margin of error. Obviously the alternative hypothesis is that all images do not have the same distribution of colour intensities within a margin of error, meaning that we can discriminate between tissue with and without carcinoma.

Based upon the literature, it is expected that images that contain a carcinoma would have a higher percentage of red. This is because most thermographic scanners display the heat collected in a colour coordinated pattern, where red is the warmest and blue is the coolest. Therefore, in comparing images, those with a cancerous growth should show a shift in percentages as compared to images of healthy subjects. The expected shift would be more red and less blue.

Results

The first attempt to classify the pixels was done manually, by creating a binary image for each layer (colour) of the image, and looking for misclassifications. This is a manual process and would have to be done for each image prior to processing. The values obtained were:

<u>benign.jpg (non cancer subject):</u>

Threshold values
Red: 33
Green: 36
Blue: 20
Number of background pixels = 326,368
Number of foreground pixels = 1,097,882
ratio of colours in image:
% red in image = 33.04
% blue in image = 33.93
% green in image = 33.04

Bin	1	2	3	4	5	6	7	8	9	10
Red										
mean	0.0384	0.1414	0.2693	0.3449	0.4600	0.5337	0.6458	0.7585	0.8531	0.9052
var	0.0018	0.0209	0.0730	0.1197	0.2125	0.2584	0.4179	0.5761	0.7286	0.8194
Green										
mean	0.0415	0.1340	0.2535	0.3505	0.4698	0.5467	0.6453	0.7574	0.8145	no pixels
var	0.0021	0.0187	0.0652	0.1237	0.2211	0.2998	0.4170	0.5743	0.6636	
Blue										
mean	0.0314	0.1653	0.2559	0.3404	0.4501	0.5440	0.6572	0.7141	no	no
var	0.0014	0.0279	0.0663	0.1166	0.2033	0.2967	0.4327	0.5101	pixels	pixels

<u>Thermal_image_1.jpg (image with carcinoma):</u>

Threshold values
Red: 1
Green: 1
Blue: 1
Number of background pixels: 19,740
Number of foreground pixels: 114,510
ratios of colours in image
% red in image = 35.02
% blue in image = 31.52
% green in image = 33.45

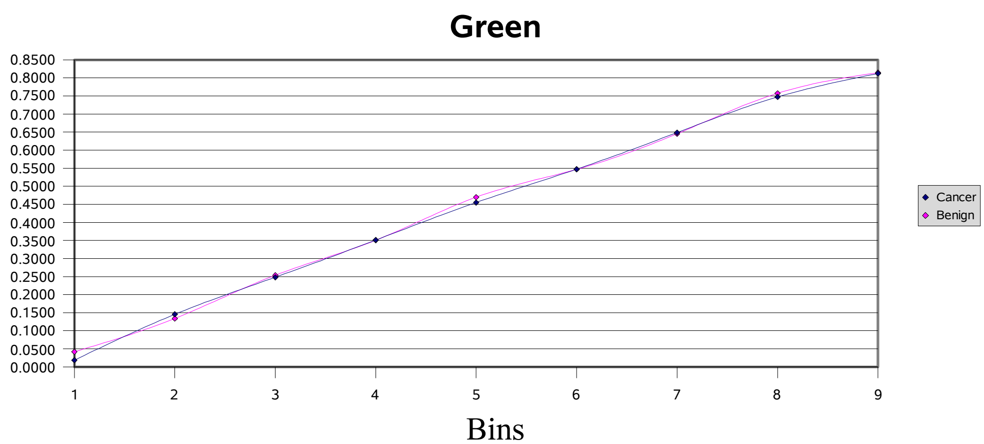
Bin	1	2	3	4	5	6	7	8	9	10
Red										
mean	0.0227	0.1493	0.2430	0.3490	0.4484	0.5478	0.6620	0.7508	0.8505	0.9263

var 0.0012 0.0231 0.0599 0.1226 0.2019 0.3009 0.4390 0.5646 0.7243 0.8585

Green
mean 0.0186 0.1454 0.2482 0.3509 0.4551 0.5472 0.6486 0.7473 0.8122 no
var 0.0009 0.0219 0.0625 0.1239 0.2080 0.3002 0.4215 0.5593 0.6598 pixels

Blue
mean 0.0209 0.1448 0.2468 0.3530 0.4455 0.5486 0.6534 0.7505 0.8518 0.9376
var 0.0010 0.0218 0.0618 0.1254 0.1994 0.3018 0.4278 0.5640 0.7264 0.8797

Graph of mean value in each bin, for both Cancer and Benign



Using the above graph of mean value per bin versus bin number, there appears to be very little difference between the two plots. It is not worthwhile to proceed with any statistical analysis on these data. The quantity of pixels per bin should hold more in the red bins for the cancer image than for the benign image, if theory holds true. The reverse should be true for the blue pixel count.

The more automated classification procedure was employed, that looked at the intensity of all three colours of each pixel as well as the total intensity. Following are the results from this more general approach.

Benign:
Number of pixels in:
background = 326368
foreground = 1097882

Percentage of image that has pixels of each colour:
red = 33.04%
green = 33.04%
blue = 33.93%

Thermal_image_1:
Number of pixels in:
background = 19740
foreground = 114510

Percentage of image that has pixels of each colour:
red = 35.02%
green = 33.45%
blue = 31.54%

The results are almost the same as the manual process, which is good that there is agreement among the methods. It gives more confidence in both the results and the methodologies.

The distribution of colours in the image are is very healthy in the first image, where there is no carcinoma. They are all close to 0.33 (percent red in image = 0.3304, percent blue in image = 0.3393, percent green in image = 0.3304). But in the second image, the one with a tumor, the red has a larger portion, and the blue a lower one (percent red in image = 0.3502, percent blue in image = 0.3152, percent green in image = 0.3345). This may seem like a small change, but it is really a large number of pixels that have become warmer (changed from blue to red). More red means more heat, and less blue means less of the lower heat. This supports the theory, as this is an indication that there is more heat in the malignant tissue image.

Unfortunately, the sample size is too small to draw any conclusive statements, other than the results of statistical analysis of 2 images are in support of the theory and that further testing needs to be done.

The pixel classification method is sufficiently complicated that it could not be done by an average radiology technician. Therefore a more automated technique is needed. The second method which compares the intensity of levels of all of the colours of a pixel, as well as a total intensity is the beginning of an automated process.

Had there been more time for this research (maybe a year instead of 3 weeks), then a neural network classifier would have been an appropriate way of classifying the pixels into foreground or background. Multi dimensional classification would be ideal, where inputs would be things like the intensity levels of the 3 colours of the pixel itself, and those of its neighbors. Other criterai, such as gradient analysis (edge detection), pixels a little farther away (bigger neighborhood), and other information could also be included to help classify a pixel. As in regression, the training process will let the system apply weights to each variable as needed to maximize accuracy, as defined by the objective function. This is a combination of neighborhood non linear filtering and the mathematical threshold analysis just performed. Additionally, it would process faster than the math and logic system that was just done. Once the neural network is trained, then it is just a matter of providing the inputs and obtaining the output.

Conclusions

Not every paper that has been written on this subject has been reviewed, but those that were read were consistent in how the data was processed. Many of the thermographs were read by a human, as a radiologist reads an X-ray. No where was it found where statistical analysis of the pixel intensity was performed. The literature supports the theory that a change in heat produced is indicative of some physiological problem. One article, mentioned earlier, stated that the average image containing a carcinoma was almost a degree warmer. It might not always be cancer but it does indicates the need for further attention.

The statistical method employed here is purely statistical and does not require a trained person to 'read' the image. Our method looks for an increase in heat produced and compares the amount of heat in the image under analysis to a benchmark value. This increase in heat is demostrated by a shift in the number of pixels from blue to red. This can be calculated by either the number of pixels or by the percentage of the image. A healthy person seems to be fairly balanced between the three primary colours, while the person with cancer appears to have a higher percentage of red and a lower percentage of blue. While the sample size was far too small, the results and the methodologies combined with the theoretical support demands further research in this area, with a larger sample size.

Future

The literature boasts almost one million instances of this methodology, with varying degrees of success. Some of the papers that were reviewed show methodology that is less than optimal (in the author's mind) and leaves room for improvement. Proper research methodologies need to be employed starting from research design and through the entire experiment. Combining improved experimental design with a larger sample size would provide a worthwhile research project, with clearly beneficial results for human health and the body of knowledge. An automated classification program as outlined above would increase accuracy. The balance of the code could be easily programmed in any of several computer languages providing a robust system for fast detection of cancer or other health concern.

References

American Cancer Society, 2000, *Breast Cancer Guidelines and Statistics*

Alamu, W., 2002, *A Review of Breast Thermography*, The International Academy of Clinical Thermology, The Internet

Darby, S. & de Gonzalez, A., 2004, *Risk of cancer from diagnostic X-rays: estimates for the UK and 14 other countries,* Lancet, 363:345-351

Epstein, S. YEAR, *the Politics of Cancer*, 539

Gotzsche,P. & Olsen, A., 2000, *Is screening for breast cancer with mammography justifiable?,* Lancet, 355:129-134

Gros, C., Gautherie, M., 1980, *Breast Thermography and Cancer Risk Prediction*. Cancer 45:51-56 (in Alamu)

Jones, B. & Plassmann, P., 2002, *Digital Infrared Thermal Imaging of Human Skin*, IEEE Engineering in Medicine and Biology, 21(6): 80-5

Kieth, L., Oleszczuk, J., & Laguens, M., 2001, *Circadian Rhythm Chaos: A New Breast Cancer Marker*, International Journal Fertility and Women's Medicine, 6(5): 238-48

Ng, E., & Sudharsan, N., 2004, *Computer simulation in conjunction with medical thermography as an adjunct tool for early detection of breast cancer;* BMC Cancer. 2004; 4: 17.

Ng, E., Fok, S., Peh, Y., Ng, F., & Sim, L., 2002, *Computerized Detection of Breast Cancer with Artificial Intelligence and Thermograms*, Journal of Medical Engineering and Technology 26(4): 152-157

Ng, E., Ung, L., Ng, F., & Sim, l., 2001, *Statistical Analysis of Healthy and Malignant Breast Thermography*, Journal of Medical Engineering and Technology, 25(6): 253-263

Nyirjesy, I, 1982, *Breast Thermography*, Clinical Obstetrics and Gynecology, 25(2):401-8

Qi, H., Snyder, W., Head, J., & Elliott, R., 2000, *Detecting Breast Cancer from Infrared Images by Asymmetry Analysis* Proceedings of the 22nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2: 1227-1228, Chicago

Simpson H., Pauson AW, Griffiths K et al. *Genesis of breast cancer is in the premenopause*. Lancet 1988; XX: 74-76. in Kieth et al

Simpson, H., McArdle, C., Pauson, A., Hume, P., Turkes, A., & Griffiths, K.; 1995, *A Non-invasive Test for the Pre-cancerous Breast,* European Journal of Cancer 31A(11) pp. 1768-1772

Wright T, McGechan A., 2003, *Breast cancer : new technologies for risk assessment and diagnosis.* Mol Diagn. 2003 Mar;7(1):49-55.