

Hyperspectral Image Processing for Detection and Grading of Skin Erythema

Ali Madooei, Ramy Mohammed Abdlaty, Lilian S Doerwald-Munoz, Joseph S Hayward, Mark S Drew, Qiyin Fang, Josiane Zerubia

▶ **To cite this version:**

Ali Madooei, Ramy Mohammed Abdlaty, Lilian S Doerwald-Munoz, Joseph S Hayward, Mark S Drew, et al.. Hyperspectral Image Processing for Detection and Grading of Skin Erythema. SPIE Medical Imaging, Feb 2017, Orlando, Florida, United States. hal-01408854v2

HAL Id: hal-01408854

https://inria.hal.science/hal-01408854v2

Submitted on 13 Dec 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Hyperspectral Image Processing for Detection and Grading of Skin Erythema

Ali Madooei[a], Ramy Mohammed Abdlaty[b], Lilian Doerwald-Munoz[c], Joseph Hayward[c], Mark S. Drew[a], Qiyin Fang[b], and Josiane Zerubia[d]

[a]School of Computing Science, Simon Fraser University, Vancouver, Canada
[b]Department of Engineering Physics, McMaster University, Hamilton, Canada
[c]Juravinski Cancer Centre, Hamilton Health Sciences, Hamilton, Canada
[d]Université Côte d'Azur, Inria, France

ABSTRACT

Visual assessment is the most common clinical investigation of skin reactions in radiotherapy. Due to the subjective nature of this method, additional noninvasive techniques are needed for more accurate evaluation. Our goal is to evaluate the effectiveness of hyperspectral image analysis for that purpose. In this pilot study, we focused on detection and grading of skin Erythema. This paper reports our proposed processing pipeline and experimental findings. Experiments have been performed to demonstrate the efficacy of the proposed approach for (1) reproducing clinical assessments, and (2) outperforming RGB imaging data.

Keywords: Hyperspectral Imaging, Skin, Erythema, Detection and Grading, Linear Discriminant Analysis

1. INTRODUCTION

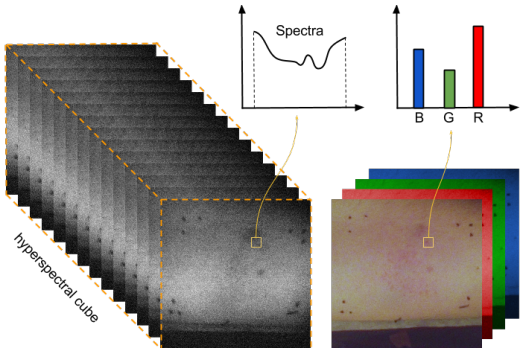


Figure 1. A schematic representation of hyperspectral vs. RGB image data. The image shows artificially induced erythema over the inside of the forearm of a volunteer (see §2.2).

Acute skin erythema is a common side effect in patients undergoing radiotherapy treatment.[1] It displays itself as an increase in skin redness and irritation. Erythema has been reported[2,3] to correlate to individual patient response to radiation and therefore may be useful to guide and modify courses of treatment in a timely manner. Currently, upon visual examination, a qualitative score can be assigned to characterize the severity of erythema, which then may be used for assessing radiation response. Although the presence or absence of erythema is relatively easily observed (since our eyes are excellent sensing devices for threshold events), characterizing the degree of erythema is difficult, imprecise and subjective (just as our sense mechanisms are inadequate in delineating graded visual and temporal differences). This highlights the need for the development of noninvasive and rapid techniques for objective and quantitative assessment of erythema.

Further author information: A.M.: amadooei@cs.sfu.ca, R.M.A: mohammrm@mcmaster.ca, L.D.M: ldoerwal@hhsc.ca, J.H.: jhayward@hhsc.ca, M.S.D.: mark@cs.sfu.ca, Q.F.: qiyin.fang@mcmaster.ca, J.Z.: josiane.zerubia@inria.fr

Previous studies have mainly focused on tissue reflectance spectroscopy[4–8] or imaging photography.[9–13] The former retrieves spectral information from point measurements* while the latter is obtained with conventional RGB colour cameras. RGB cameras divide the light spectrum into broad overlapping red, green, and blue image slices that when combined seem realistic to the eye. Photography has the advantage of offering spatial information† but this comes at the cost of losing much of spectral information. We use hyperspectral imaging (HSI) which provides both spatial and spectral representation of the affected area. Here, a hyperspectral camera effectively divides the spectrum into very many thin image slices (the actual number depending on the camera and application: Fig.1). This fine-grained slicing reveals spectral structure that may not be evident to the eye or to an RGB camera but can provide a rich set of information for image processing. HSI techniques have widely demonstrated their usefulness in different areas of medical imaging research during the last decade.[14] As an emerging imaging modality for medical applications, the combination of HSI devices with adequate image processing techniques offers the perfect landscape for developing new methods for noninvasive disease monitoring and diagnosis. Nevertheless, the instrumentation and subsequent image analysis must be thoughtfully considered for adequately revealing the useful and relevant information.

The purpose of the present study was to investigate the possibility of monitoring the degree of erythema using HSI data. This is the first step towards building a more reliable computer-aided imaging system for monitoring radiation-induced skin reactions. Such a system can provide clinicians with an objective second opinion. This paper describes our proposed processing pipeline and experimental findings. We evaluate the system against the clinical assessment of an experienced clinician. We also compare the performance to that of using digital photography (instead of HSI). The results from this preliminary study are encouraging and indicate that hyperspectral image data do contain relevant information, and indeed outperform imaging photography.

2. MATERIAL AND METHODS

2.1 Hyperspectral Imaging System

Our collaborators at McMaster University have been investigating the development of clinically suitable optical imaging modalities including hyperspectral imaging (e.g.[15]). For this study, an HSI system was developed which uses acousto-optical tunable filter (AOTF) technology. "An AOTF consists of a crystal in which radio frequency acoustic waves are used to separate a single wavelength of light from a broadband source".[16] The developed AOTF-HSI system operates in the spectral range of 450-800 nm. The image detector incorporated in the system is a monochromatic Complementary Metal Oxide Semiconductor (CMOS) camera with active area of 11.24 mm^2 (2048×2048 pixels). The CMOS sensor has the advantages of compactness and relative cheap cost with respect to corresponding charge-coupled device (CCD) sensors (with the same sensor size). The system uses three distinct halogen sources for illumination with total of 550 watts.

2.2 Clinical Data Collection

Radiation induced erythema is a result of inflammation and dilation of skin capillaries.[2] It develops gradually over the course of treatment and usually peaks during the 5^{th} week of therapy. For the purpose of controlled experiments, it is often more convenient to artificially induce erythema using some sort of skin irritant (e.g. lidocaine, UV light, surfactants such as Sodium Lauryl Sulfate or skin abrasion).[17]

In this study, healthy adult volunteers were recruited to participate in a controlled experiment. A rectangular area was marked at the corners with blue or green marker on the inside of the forearm of the volunteer (Fig.1). This area was chosen because it was a convenient area of flat skin easy to reposition and image repeatedly. The rectangular area would be the region of interest (ROI) for this experiment and was imaged with the HSI system and RGB camera.‡ Baseline image of the ROI were taken before inducing erythema. Following base-line

*Point measurements use probes that come in contact with the skin surface, deliver light to the point of contact and collect the remitted light.

†The spatial information is useful because although the prescription radiation dose is the same for the entire affected area, electron scatter can result in non-uniform development of erythema.

‡The RGB camera was oriented in the same direction and under same lighting condition as the HSI camera so that HSI and RGB images were comparable to facilitate feature comparison between both image modalities.

imaging, erythema was artificially induced in the ROI: volunteers were asked to use a plastic bar to strike the skin in rapid succession for approximately 3 minutes until the skin would become bright red.[§] Immediately after erythema was induced, the ROI was imaged again with both technologies at regular intervals until erythema was not visible anymore.

For clinical assessment of erythema, a new scoring system (Table-1) was designed for this experiment to capture intermediate stages of skin erythema reactions. This new erythema score focuses on the degree of redness of the skin (pink to red) and the clarity of the borders defining the erythemic areas. During the experiment, the clinician would score erythema of the ROI. After the experiment was completed, the clinician annotated RGB images by contouring erythema areas and assigning erythema scores to correspond with the ones previously recorded during clinical assessment. These annotations were used for image classification (see §2.4 & §3).

Table 1. Erythema grading schema.

Condition	Score	Description
Very faint erythema	1	Appears as a slight increase in redness (very light pink) of the skin from baseline.
Faint erythema	2	Skin reaction is more apparent. Borders are more clearly defined.
Bright erythema	3	Erythema is clearly apparent, skin is bright pink and borders clearly defined.
Very bright erythema	4	Erythema is clearly apparent, skin is bright red and borders clearly defined.

2.3 HSI Data Correction and Pre-processing

In order to directly compare HSI spectra with each other, the encoded radiance values in the RAW data must be converted to reflectance. A comprehensive conversion must account for various imaging factors, in particular the spectrum of light source. The RAW data is converted to "relative" reflectance by dividing each image band (HSI spectrum) by the (mean) spectrum of a "white surface" that has a relatively flat spectral reflectance curve. The white surface can be included in the scene or imaged prior to each acquisition.

Accounting for lighting effects due to the angle and topography of sources of illumination is among the major challenges of our work. We have found empirically that this defect can be partly removed by subtracting one of HSI image bands from the entire HSI image cube. We have subtracted the image band corresponding to AOTF tuning frequency of 141 MHz which captures information at wavelength 446 nm. This wavelength corresponds to the lower end of the spectrum and particularly shows a very low output in the spectrum of our light source. The corresponding image band mainly contains the pattern of (uneven) illumination, camera's thermal noise ("dark current") and sensor's fixed pattern noise ("hot pixels"). Thus, the subtraction also partially corrects the RAW data for dark current effect and noise due to possible faulty or "electronic bias" of the sensor.

Hyperspectral images suffer from instrument noise. This defect can be partly removed by using smoothing techniques and many different algorithms are used in smoothing. We use a Wiener filter,[18] a low-pass image smoothing technique which smooths an image adaptively, tailoring itself to the local image variance.

$$J(x) = \mu + \frac{\sigma^2 - v^2}{\sigma^2}\left(I(x) - \mu\right) \tag{1}$$

where J is the filtered image, I is the original input image, x indexes over pixels, v is the noise variance and, μ and σ^2 are locally computed (refer to[18] for more information).

2.4 Classification

We use linear discriminant analysis (LDA) for classification.[19] Consider the general setting of a discriminant algorithm for solving classification problem given labelled observations $(\mathbf{x}, y)$ where $\mathbf{x} \in R^p$ is a p-dimensional input vector and y is one of the $1 \ldots K$ discrete classes. A discriminant function can be defined as

$$h(\mathbf{x}) = \arg\max_k P(y = k|\mathbf{x}) \tag{2}$$

[§]No long-term harm to volunteers resulted from artificially inducing erythema of the skin in this manner other than transient redness of the skin and some minor bruising that resolved without treatment.

where the class posterior is defined according to the Bayes rule:

$$P(y=k|\mathbf{x}) = \frac{P(\mathbf{x}|y=k)P(y=k)}{P(\mathbf{x})} \tag{3}$$

The prior probability $P(y=k)$ is empirically set to n_k/n where n represents number of observations (hence n_k is the number of observation for class k). The $P(\mathbf{x})$ is a normalization constant and it is defined as the summation over k of $P(\mathbf{x}|y=k)P(y=k)$.

LDA models the conditional class probability $P(\mathbf{x}|y=k)$ using a multivariate normal density function:

$$f(\mathbf{x}|\mu_k, \mathbf{\Sigma}_k) = \frac{\exp\{-\frac{1}{2}(\mathbf{x}-\mu_k)^T \mathbf{\Sigma}_k^{-1}(\mathbf{x}-\mu_k)\}}{\sqrt{2\pi|\mathbf{\Sigma}_k|}} \tag{4}$$

where μ_k and $\mathbf{\Sigma}_k$ are class mean and covariances; $|\cdot|$ and $\cdot^{-1}$ are matrix determinant and inverse, respectively. LDA further assumes the classes have a common covariance matrix $\mathbf{\Sigma}_k = \mathbf{\Sigma}\ \forall k$ which leads to linear decision boundaries. Under these assumptions, LDA infers the mean $\hat{\mu}_k$ and covariance $\hat{\mathbf{\Sigma}}$ parameters from training data:

$$\hat{\mu}_k = \frac{1}{n_k} \sum_{y_i=k} \mathbf{x}_i \tag{5}$$

$$\hat{\mathbf{\Sigma}} = \frac{1}{n-K} \sum_{k=1}^{K} \left(\sum_{y_i=k} (\mathbf{x}_i - \hat{\mu}_k)(\mathbf{x}_i - \hat{\mu}_k)^T \right) \tag{6}$$

2.4.1 Learning with noisy labels

One of the challenges of this work is dealing with noisy labels due to the imperfect registration between the RGB image (which was used by the clinician for annotation) and the hyperspectral image cube, as well as small misalignments between the hyperspectral image bands (due to e.g. subject movement during imaging). To compensate for these defects, we construct a weighted LDA classifier where each data point is weighted with weights w_i. The parameters of the model are inferred by modifying eq.5 and eq.6:

$$\hat{\mu}_k = \frac{\sum_{y_i=k} w_i \mathbf{x}_i}{\sum_{y_i=k} w_i} \tag{7}$$

$$\hat{\mathbf{\Sigma}} = \frac{\sum_{k=1}^{K} \left(\sum_{y_i=k} w_i (\mathbf{x}_i - \hat{\mu}_k)(\mathbf{x}_i - \hat{\mu}_k)^T \right)}{1 - \sum_{k=1}^{K} \left(\sum_{y_i=k} w_i^2 / \sum_{y_i=k} w_i \right)} \tag{8}$$

The weights vector $\mathbf{w}$ is learned using a boosting method similar to that in.[20] Alg-1 summarizes this method.

Algorithm 1 – Boosting LDA

1: Given N samples of training set $(\mathbf{x}_i, y_i)$
2: Initialize $w_i = 1/N$
3: Set λ [an empirical parameter used to tune the magnitude of the penalization for each iteration (e.g., $\lambda = 1/2$)].
4: **for** $m = 1$ to M **do**
5: Obtain $(\hat{\mu}_k, \mathbf{\Sigma})$ according to eq.7 and eq.8
6: Predict the labels $\hat{y}_i = h(\mathbf{x}_i|\mu_k, \mathbf{\Sigma})$
7: Compute the error $\varepsilon = \sum_i (w_i \times I(y_i \neq \hat{y}_i)) / \sum_i (w_i)$
8: Compute $\alpha = \lambda \times \log\{(1-\varepsilon)/\varepsilon\}$
9: Update the weight $w_i = w_i \times \exp\{\alpha \times I(y_i \neq \hat{y}_i)\}$
10: Normalize $\mathbf{w} \leftarrow \mathbf{w}/||\mathbf{w}||$
11: **end for**

The idea of introducing sample weights is to give more influence to those samples in the training set that are labelled correctly.

2.4.2 Dealing with noisy measurements

In our imaging system, image bands corresponding to the lower end of the spectrum exhibit very low signal to noise ratio as the result of low sensitivity of the sensor and low output of the light source in that region of the spectrum. We have found empirically that removing these bands will improve classification performance. We impose the problem of systematically choosing a set of bands to be removed as a problem of matrix "Column Subset Selection". Suppose $\mathbf{X} \subset R^p$ is an $(n \times p)$ dimensional data set representing n pixels from training data. The p columns are representing spectral bands and, the problem is the selection of a subset of columns from $\mathbf{X}$ that best approximate its span. We use a simple technique based on "matrix low-rank representation" to remove the columns (bands) that contribute the least to the "information" contained within the data set $\mathbf{X}$. We quantify this information using the Frobenius (element-wise) matrix norm: $\|\mathbf{X}\| = \sqrt{\sum_{i,j}(\mathbf{X}_{i,j})^2}$.

We first compute the sum of squares of each column of data matrix and sort these in descending order. We then incrementally add these values to the norm $\|\hat{\mathbf{X}}\|$ while $\left(\|\hat{\mathbf{X}}\| / \|\mathbf{X}\|\right) \geqslant p$ where $p \in [0, 1]$ is a threshold value. Retaining 99% of the information (p = 0.99) reduces the number of columns to 79, removing 10 spectral bands (Fig-2).

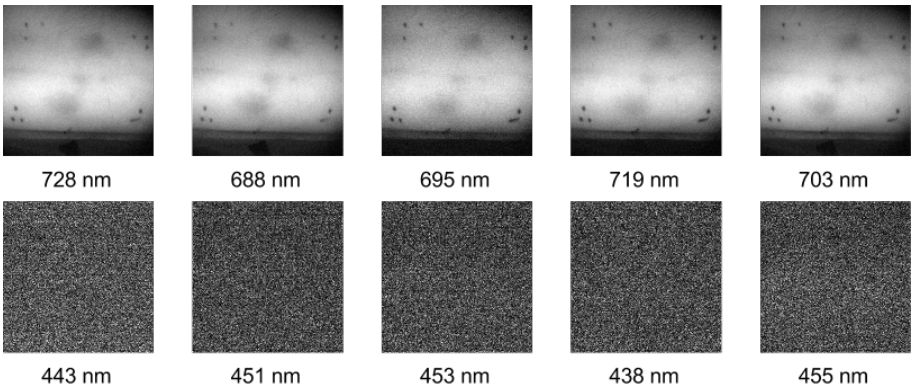


Figure 2. Top 5 bands contributing highest (top row) and least (bottom row) to the HSI information (best viewed in PDF).

3. EXPERIMENT AND RESULTS

At this stage in our study, three volunteers have been recruited for controlled experiments and four experiments have been conducted. In total, we obtained 12 set of HSI and RGB images of which 2 were used for training and the remaining 10 for testing the proposed system. Since we perform pixel-level classification of erythema, these few images account to more than 50 million data samples (each image is 2048×2048 pixels). In fact, we performed down-sampling (each image was scaled by a factor of 0.4) to reduce the data set. Table-2 summarizes the results. We report accuracy, precision, recall and f-measure.[21] For comparison, we also report the classification results using RGB data. Note that HSI outperforms RGB in every aspect, by a large margin.

Table 2. Results for pixel-level classification of erythema.

Method	Accuracy	Precision	Recall	f-score
HSI	**87.26**	**76.18**	**78.86**	**76.11**
RGB	70.24	71.25	55.96	62.07

We also combined the problem of erythema detection and grading into a multi-class classification problem where each pixel is classified as one of 4 erythema classes (see Table-1), or non-erythema class. The results are encouraging as it is indicative that (1) different classes of erythema are linearly separable in the hyperspectral domain, and (2) the spectral information beyond the visible spectrum (especially middle and near infrared) is also useful for this task.

4. CONCLUSION

We have shown that hyperspectral image processing can be successfully employed for detection and grading of erythema. HSI is much more powerful than traditional photography. It can provide a rich set of information to enable and accelerate the automation of tasks in Medical Image Analysis.

In future, we want to extend the technique to further detect other skin responses to radiation (such as dry/moist desquamation, skin necrosis, etc.) and also to experiment with real patients undergoing radiotherapy. Our ultimate objective is to build a system for monitoring radiation response in individuals using HSI technology and image processing.

5. ACKNOWLEDGEMENT

The work of the first author was partly supported by MITACS-Inria Globalink Research Award (IT05363) during his stay in Ayin Team at Inria Sophia-Antipolis Méditerranée.

REFERENCES

[1] J. Eisenberg, "Comparative effectiveness and safety of radiotherapy treatments for head and neck cancer," in [*Comparative Effectiveness Review Summary Guides for Clinicians*], Agency for Healthcare Research and Quality (2007).

[2] Denham, J., Hamilton, C., Simpson, S., Ostwald, P., O'Brien, M., Kron, T., Joseph, D., and Dear, K., "Factors influencing the degree of erythematous skin reactions in humans," *Radiotherapy & Oncology* **36**(2), 107–120 (1995).

[3] Turesson, I., Nyman, J., Holmberg, E., and Oden, A., "Prognostic factors for acute and late skin reactions in radiotheraphy patients," *J. of Radiation Oncology, Biology & Physics* **36**(5), 1065–1075 (1996).

[4] Riordan, B., Sprigle, S., and Linden, M., "Testing the validity of erythema detection algorithms," *J. of rehabilitation research and development* **38**(1) (2001).

[5] Wengstrom, Y., Forsberg, C., Naslund, I., and Bergh, J., "Quantitative assessment of skin erythema due to radiotherapy – evaluation of different measurements," *Radiotherapy & Oncology* **72**(2), 191–197 (2004).

[6] Nyström, J., Geladi, P., Lindholm-Sethson, B., Rattfelt, J., Svensk, A., and Franzen, L., "Objective measurements of radiotherapy-induced erythema," *Skin Research and Technology* **10**(4), 242–250 (2004).

[7] Myers, D., Cooper, C., Beilman, G., Mowlem, J., Anderson, L., Seifert, R., and Ortner, J., "A wide gap second derivative NIR spectroscopic method for measuring tissue hemoglobin oxygen saturation," in [*Oxygen Transport to Tissue XXVII*], 217–222, Springer (2006).

[8] Stamatas, G., Zmudzka, B., Kollias, N., and Beer, J., "In vivo measurement of skin erythema and pigmentation: new means of implementation of diffuse reflectance spectroscopy with a commercial instrument," *British J. of Dermatology* **159**(3), 683–690 (2008-09).

[9] Nischik, M. and Forster, C., "Analysis of skin erythema using true-color images," *Medical Imaging, IEEE Transactions on* **16**(6), 711–716 (1997).

[10] Setaro, M. and Sparavigna, A., "Quantification of erythema using digital camera and computer-based colour image analysis: a multicentre study," *Skin Research and Technology* **8**(2), 84–88 (2002).

[11] Yamamoto, T., Takiwaki, H., Arase, S., and Ohshima, H., "Derivation and clinical application of special imaging by means of digital cameras and ImageJ freeware for quantification of erythema and pigmentation," *Skin Research and Technology* **14**(1), 26–34 (2008).

[12] Lu, J., Manton, J., Kazmierczak, E., and Sinclair, R., "Erythema detection in digital skin images," in [*2010 17th IEEE International Conference on Image Processing (ICIP)*], 2545–2548 (2010).

[13] Matsubara, H., Matsufuji, N., Tsuji, H., Yamamoto, N., Karasawa, K., Nakajima, M., Takahashi, W., and Karube, M., "Objective assessment in digital images of skin erythema caused by radiotherapy," *Medical Physics* **42**(9), 5568–5577 (2015).

[14] Lu, G. and Fei, B., "Medical hyperspectral imaging: a review," *J. of Biomedical Optics* **19**(1) (2014).

[15] Yuan, Y., Hwang, J., Krishnamoorthy, M., Ye, K., Zhang, Y., Ning, J., Wang, R., Deen, M., and Fang, Q., "High-throughput acousto-optic-tunable-filter-based time-resolved fluorescence spectrometer for optical biopsy," *Optics Letters* **34**(7), 1132–1134 (2009).

[16] Gat, N., "Imaging spectroscopy using tunable filters: a review," *Proc. SPIE, Wavelet Applications VII* **4056**(50), 50–64 (2000).

[17] Bodekær, M., Philipsen, P. A., Karlsmark, T., and Wulf, H., "Good agreement between minimal erythema dose test reactions and objective measurements: an in vivo study of human skin," *Photodermatology, Photoimmunology & Photomedicine* **29**(4), 190–195 (2013).

[18] Jain, A., [*Fundamentals of digital image processing*], Prentice Hall (1989).

[19] Hastie, T., Tibshirani, R., and Friedman, J., [*The Elements of Statistical Learning*], Springer Series in Statistics, Springer New York (2009).

[20] Wang, B. and Japkowicz, N., "Boosting support vector machines for imbalanced data sets," *Knowledge and Information Systems* **25**(1), 1–20 (2009).

[21] Demšar, J., "Statistical comparisons of classifiers over multiple data sets," *J. of Machine Learning Research* **7**, 1–30 (2006).