Optimal Feature Selection Applied to Multispectral Fluorescence Imaging

Tobias C. Wood[1], Surapa Thiemjarus[1], Kevin R. Koh[1], Daniel S. Elson[1,2], and Guang-Zhong Yang[1]

[1] Institute of Biomedical Engineering, Imperial College London
[2] Department of Biosurgery and Surgical Technology, Imperial College London
g.z.yang@imperial.ac.uk

Abstract. Recent rapid developments in multi-modal optical imaging have created a significant clinical demand for its *in vivo - in situ* application. This offers the potential for real-time tissue characterization, functional assessment, and intra-operative guidance. One of the key requirements for *in vivo* consideration is to minimise the acquisition window to avoid tissue motion and deformation, whilst making the best use of the available photons to account for correlation or redundancy between different dimensions. The purpose of this paper is to propose a feature selection framework to identify the best combination of features for discriminating between different tissue classes such that redundant or irrelevant information can be avoided during data acquisition. The method is based on a Bayesian framework for feature selection by using the receiver operating characteristic curves to determine the most pertinent data to capture. This represents a general technique that can be applied to different multi-modal imaging modalities and initial results derived from phantom and *ex vivo* tissue experiments demonstrate the potential clinical value of the technique.

Keywords: Feature Selection, Fluorescence Imaging, Multispectral Imaging, BFFS, Receiver Operating Characteristic.

1 Introduction

Fluorescence spectroscopy is an established technique for investigating tissue properties. Normal tissue contains several types of endogenous fluorophore, including collagen, elastin, porphyrins, NADH and FAD. Their relative abundances may be detected through multispectral fluorescence methods to distinguish different disease states [1,2]. In addition, quantitative optical molecular *in vivo* imaging using fluorescently labelled tumours can be achieved, allowing preclinical studies using small animal models [3]. Photosensitising chemicals can also be used to synthesise fluorescent compounds directly within cells, such as PPIX derived from 5-ALA, allowing cancerous tissue to be characterised in human patients [4]. Multispectral fluorescence measurements are also being used to study protein-protein interactions based on Föster Resonance Energy Transfer (FRET)

[3]. Fluorescence light can hence provide a rich source of information about tissue composition, structure and even cell metabolism, although the complexity of the multispectral signal requires significant pre-processing and interpretation.

Thus far, multispectral fluorescence has mainly been acquired from single points (non imaging), but more recently there has been a trend towards acquiring spectral images due to the development of CCDs and tuneable filters. These systems often acquire a reduced number of spectral points due to the limited number of photons per pixel of the detection system. The success of *in vivo* tests, administered *in situ* at the patient location, relies on maximising sensitivity and specificity, whilst at the same time procedure times need to be minimised to increase throughput and reduce patient discomfort.

Multispectral selection and classification is important in making best use of the available photons and reducing acquisition times whilst accounting for correlation or redundancy between different dimensions. Post-processing techniques such as Isomap [5] and Spectral Unmixing [3] are useful techniques for visualising datasets but require the capture of a complete spectrum. Previous attempts at quantifying excitation and emission wavelengths that contain the most diagnostic information have been primitive in two respects. In general only a small number of excitation wavelengths have been available due to hardware limitations, although more advanced illumination sources have been reported [6]. The resulting data has then only been analysed using linear techniques such as Principal Component Analysis (PCA) [7] or auto-correlation vectors [8] in order to linearly rank the best spectroscopic information. However combinations of wavelengths are ignored, thus reducing the possible discriminating power.

Our system, still in development, uses a broadband supercontinuum laser source and a movable slit to provide high resolution spectral selection in combination with a CCD allowing an entire image to be captured simultaneously. It has the significant advantage of being able to automatically select both the central excitation wavelength and the spectral bandwidth freely so that the optimal combinations of wavelengths may be selected without requiring the modification of filters or light sources. We then propose a feature selection framework based on Bayesian principles to identify the best combinations of wavelengths for discriminating between different tissue classes. The method was originally developed for use in body sensor networks, where there is also emphasis on recording only essential data in order to satisfy the resource constraints of miniaturised sensor nodes [9,10].

2   Methods

2.1   Fluorescence Imaging

The fluorescence imaging system used is shown in Fig. 1(a) and has been previously described [5,11]. With this system, samples are excited using a supercontinuum laser source that emits broadband (450-800 nm) pulsed light. Narrow portions of this spectrum are then selectively chosen by dispersing the beam with a prism and filtering undesired wavelengths using a slit. The remaining
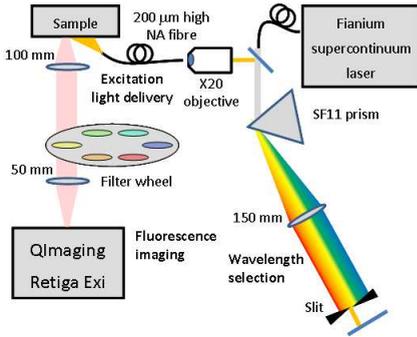
(a) Experimental Setup
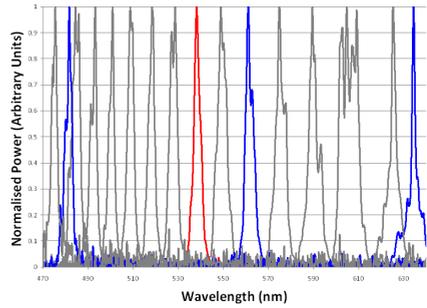
(b) Resulting Spectra

Fig. 1. The optical setup of the imaging system and resulting spectra generated. The supercontinuum generated by the laser is spatially separated by a prism so that narrow-band portions can be selected by moving the slit. The best imaging band is highlighted in red, the best combination of three in blue.

spectrum is then focused onto the sample and the resulting fluorescence imaged onto a CCD camera through a long-pass filter to block reflected excitation light.The integration period of the CCD was set to 0.8 seconds. Combined with the positioning time of the slit the total acquisition time was approximately 20 seconds. This scales linearly with the number of images captured.

The samples for this study were collagen gel tissue phantoms designed to simulate human epithelial tissue. The gels were contained in cylindrical wells with a diameter of 15 mm and were 3 mm thick. Fluorophores were added to the collagen with a concentration of 3.0 μM. The fluorophores used were Coumarin 334 (C334), Cresyl Violet (CV), Flavin Adenine Dinucleotide (FAD), Fluorescein (FL), Rhodamine B (RB), Rhodamine 6G (R6G) and Protoporphyrin IX (PPIX). Combinations of fluorophores were also used. Images of each well were captured separately and combined to form a single composite image at each wavelength. Several bright-field images were recorded at each excitation wavelength (listed in Table 1) and averaged to reduce noise. Individual pixels where then normalised to these to remove both the spatial variation in illumination across the well and the spectral power variation of the laser across wavelengths.

Table 1. Central wavelength of the excitation spectra used as features (bandwidths varied from 2 - 9 nm). The spectra are plotted in Fig. 1(b). An additional feature used was the total integrated intensity of the spectrum of each pixel.

Feature Number	1	2	3	4	5	6	7	8
Wavelength (nm)	476	482	486	493	501	510	519	529
Feature Number	9	10	11	12	13	14	15	16
Wavelength (nm)	538	549	561	575	589	605	625	634

Pixels were then normalised so their total integrated intensity across all wavelengths was unity in order to correct for variations in fluorophore concentration. The normalised 16-point excitation spectra and total integrated intensity at each pixel comprised the complete feature set.

2.2   Feature Selection

For optimal feature selection, a multi-objective Bayesian Framework for Feature Selection (BFFS) combined with a Pruned-Tree search algorithm was used [12]. BFFS is a filter-based machine learning technique for reducing model complexity to avoid over-fitting problems [10]. It selects combinations of features based on the Area Under the Curve (AUC) of their combined Receiver Operating Characteristic (ROC) curve, and thus is independent of a specific classification model. Multi-Objective BFFS extends the evaluation process to include the model complexity and feature redundancy to further optimise the discriminative power of the feature sets whilst minimising the number of spectra to be used. To this end, the following objective function was used:

$$D(f_i) = -(1 - \omega_1)E_{AUC}(G^k - f_i) + \omega_1 E_{AUC}(f_i) \tag{1}$$

where f_i is a feature from the set G^k at step k and ω_1 is a weighting factor. These weighting factors are searched using a Pruned-Tree algorithm to find the optimal solution set in an efficient manner. At each step in the Pruned-Tree algorithm, the objective function is evaluated for all features in the complete set. Backwards propagation is then used to sequentially remove the feature with the lowest objective function until only the most discriminating feature is left. An ordered list of the best feature combinations and their corresponding AUC can then be produced.

Each AUC curve is generated from the proportion of pixels in the composite image that are classified as the correct fluorophore. Because each fluorophore or mixture is localised within a single well, all pixels within the well should receive a single class label. Consequently the True Positive (TP) rate is the number of pixels in one well classified as the correct fluorophore divided by the pixel count for each well. The False Positive rate is then number of pixels classified as the same fluorophore in all other wells divided by the total pixel count of those wells.

Once these combinations had been determined, a Naïve Bayesian classifier was constructed to assign class labels to new images. A Naïve Bayesian classifier is a type of probabilistic graphical model with a strong independent assumption that all the features in the model are independent of each other. The model was trained based on maximum likelihood estimation and used only a small amount of training data

3   Results

3.1   Tissue Phantom Validation

Fig. 2 illustrates an example of the intensity images captured by the optical imaging setup after individual gel images have been stitched together. Table 2
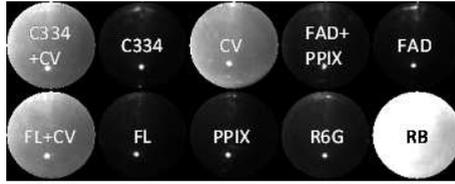
Fig. 2. A composite intensity image of the fluorophore gels captured at 475 nm. 16 such images excited at the wavelengths given in Table 1, together with the integrated total intensity image, form the complete data set for classification.

illustrates the output from the multi-objective BFFS algorithm, which shows that the expected AUC increases with the number of features used. It is evident that the increase after more than five features is not significant, showing the high level of redundancy in the data set. The actual wavelengths selected as the single best and as the best combination of three are highlighted in Fig. 1(b). The best three features are spectrally distant from each other. This is to be expected as these wavelengths are unlikely to contain much redundant information. Fig. 3 shows the AUC for classifying each individual fluorophore as the number of features is increased. With all features available seven of the ten fluorophores can be distinguished from each other almost perfectly, and the remaining three still have a high AUC of about 0.95. The AUC for each classifier only starts to drop significantly when the number of features is reduced to three, implying that the ability to discriminate between fluorophores is already high when imaging with just three wavelengths. Any additional data acquired is therefore mostly irrelevant and so does not add significant new information to that already present.

Fig. 4 provides two example ROC curves for classifying pixels as PPIX using just the single best feature (the image at 538 nm), and the best three features (482, 561 and 634 nm). With one feature the system labels many pixels as PPIX even with low thresholds, resulting in high sensitivity but also low specificity. Adding the additional features increases the specificity at all threshold values, thus the AUC rises from 0.79 to 0.92. The figures also show pictorially the classification of the fluorophores as the threshold is varied along the ROC. These indicate that the majority of false PPIX positives come from the well containing a mixture of FAD and PPIX, due to the similarity in spectra. There is a large difference between the spectra of PPIX and the other fluorophores [1], resulting in good classification accuracy.

Table 2. The expected AUC for feature sets with different numbers of features

Number of Features	1	2	3	4	5	6	7	8
Expected AUC	0.891	0.957	0.986	0.988	0.991	0.994	0.995	0.995
Number of Features	9	10	11	12	13	14+		
Expected AUC	0.996	0.996	0.996	0.997	0.997	0.998		

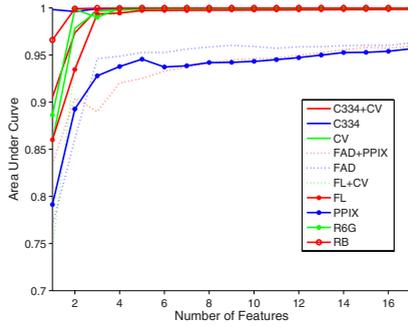**Fig. 3.** Plot of the AUC for discriminating each class from the remainder as the number of features is increased

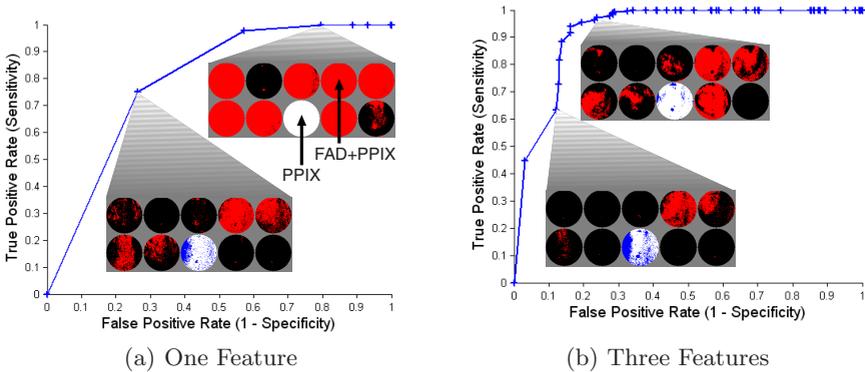(a) One Feature   (b) Three Features

Fig. 4. ROC curve for classifying PPIX against the remaining phantoms with one (a) and three (b) features and pictorial representations of the classification at selected points on the curve. True positives and negatives are indicated by white and black, false positives and negatives by red and blue, respectively.

3.2   Lung Cancer Biopsies

The proposed technique was evaluated with *ex vivo* human lung cancer biopsies. Two biopsies obtained from a single patient during a broncoscopy procedure to debulk a tumour mass are shown in Fig. 5. One biopsy was treated with 5-ALA in order to induce production of PPIX, an exogenous fluorophore. The other biopsy was untreated and so only contained endogenous fluorophores such as elastin. The excitation wavelengths of these are shorter than the those used in the imaging system, so they did not fluoresce significantly. *In vivo*, tumours are known to preferentially metabolise 5-ALA compared to healthy tissue [4]. Thus the spectral signature of PPIX can be used as a diagnostic test for the presence of cancer. Figs. 5(b) and (c) have been coloured using the probability of a pixel being classified with a Naïve Bayesian classifier as containing PPIX. The best features from section 3.1 were used to build the model, using 20% of the available
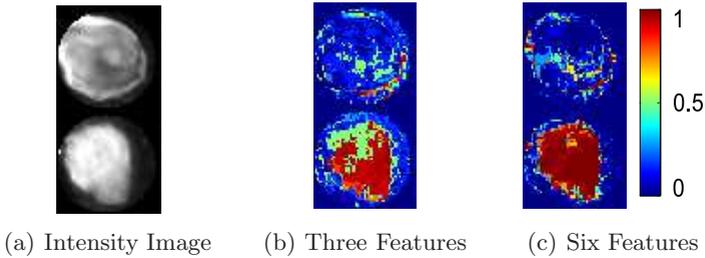
(a) Intensity Image   (b) Three Features   (c) Six Features

Fig. 5. Images of two lung cancer biopsies. The bottom biopsy has been treated with 5-ALA, whereas the top biopsy is untreated. (a) shows a typical fluorescence intensity image captured at 561 nm. (b) and (c) show false colour maps of the probability that a pixel contains PPIX, produced by Naïve Bayes classifier with three and six features respectively.

pixels. These show that the stained biopsy can be clearly distinguished from the unstained biopsy, and that the accuracy of this distinction increases with the number of features, allowing applications in fast 'red flag' screening of patients. The diagnostic protocol can also be tailored to suit specific situations. If fast procedure times are required then fewer features can be used, but when high sensitivity and specificity are paramount then a slower procedure utilising all features could be conducted.

4   Conclusion

In this paper, Multi-objective BFFS was demonstrated as a useful tool for finding the best wavelengths to use in a multispectral imaging system to distinguish between different fluorophore varieties. This information can then be fed back into the data collection method to prevent the collection of redundant data. For instance, with the fluorescence system presented in this paper, imaging at just three wavelengths, 482 nm, 561 nm and 634 nm, provides practically as much diagnostic information as when imaging with all sixteen wavelengths available, whilst decreasing acquisition times by a factor of five. The diagnostic protocol can hence be tuned to provide high throughput without significant loss of accuracy.

References

1. DaCosta, R.S., Andersson, H., Wilson, B.C.: Molecular fluorescence excitation-emission matrices relevant to tissue spectroscopy. Photochemistry and Photobiology 78(4), 384–392 (2003)
2. Schwarz, R.A., Gao, W., Daye, D., Williams, M.D., Richards-Kortum, R., Gillenwater, A.M.: Autofluorescence and diffuse reflectance spectroscopy of oral epithelial tissue using a depth-sensitive fiber-optic probe. Applied Optics 47(6), 825–834 (2008)
3. Zimmermann, T., Rietdorf, J., Pepperkok, R.: Spectral imaging and its applications in live cell microscopy. FEBS Letters 546(1), 87–92 (2003)

4. Berg, K., Selbo, P.K., Weyergang, A., Dietze, A., Prasmickaite, L., Bonsted, A., Engesaeter, B.O., Angell-Petersen, E., Warloe, T., Frandsen, N., Hogset, A.: Porphyrin-related photosensitizers for cancer imaging and therapeutic applications. Journal of Microscopy 218(2), 133–147 (2005)
5. Lekadir, K., Elson, D., Requejo-Isidro, J., Dunsby, C., McGinty, J., Galletly, N., Stamp, G., French, P., Yang, G.Z.: Tissue characterization using dimensionality reduction and fluorescence imaging. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006. LNCS, vol. 4191, pp. 586–593. Springer, Heidelberg (2006)
6. Heintzelman, D.L., Utzinger, U., Fuchs, H., Zuluaga, A., Gossage, K., Gillenwater, A.M., Jacob, R., Kemp, B., Richards-Kortum, R.R.: Optimal excitation wavelengths for in vivo detection of oral neoplasia using fluorescence spectroscopy. Photochemistry and Photobiology 72(1), 103–113 (2000)
7. Martin, S.F., Wood, A.D., McRobbie, M.M., Mazilu, M., McDonald, M.P., Ifor, W.D., Samuel, C., Herrington, S.: Fluorescence spectroscopy of an in vitro model of human cervical precancer identifies neoplastic phenotype. International Journal of Cancer 120(9), 1964–1970 (2007)
8. Zuluaga, A.F., Utzinger, U., Durkin, A., Fuchs, H., Gillenwater, A., Jacob, R., Kemp, B., Fan, J., Richards-Kortum, R.: Fluorescence excitation emission matrices of human tissue: A system for in vivo measurement and method of data analysis. Applied Spectroscopy 53, 302–311 (1999)
9. Thiemjarus, S., Lo, B.P.L., Yang, G.Z.: Feature selection for wireless sensor networks. In: First International Workshop on Wearable and Implantable Body Sensor Networks, Imperial College, London (2004)
10. Yang, G.Z., Hu, X.: Multi-sensor fusion. In: Yang, G.Z. (ed.) Body Sensor Networks, pp. 262–280. Springer, Heidelberg (2006)
11. Koh, K., Wood, T., Zhang, H., Lekadir, K., Elson, D., Yang, G.Z.: Fluorescence excitation spectroscopic imaging with a tunable light source and dimensionality reduction using fr-isomap. In: SPIE BiOS (2008)
12. Thiemjarus, S., Yang, G.Z.: An autonomic sensing framework for body sensor networks. In: Second International Conference on Body Area Networks, Florence, Italy (2007)