World Academy of Science, Engineering and Technology
International Journal of Medical and Health Sciences
Vol:2, No:8, 2008

Enhancement of m-FISH Images using Spectral Unmixing

Martin De Biasio, Raimund Leitner, Franz G. Wuertz, Sergey Verzakov and Pierre J. Elbischger

Abstract*—Breast carcinoma* is the most common form of cancer in women. Multicolour fluorescent in-situ hybridisation (m-FISH) is a common method for staging *breast carcinoma*. The interpretation of m-FISH images is complicated due to two effects: (i) Spectral overlap in the emission spectra of fluorochrome marked DNA probes and (ii) tissue autofluorescence. In this paper hyper-spectral images of m-FISH samples are used and spectral unmixing is applied to produce false colour images with higher contrast and better information content than standard RGB images. The spectral unmixing is realised by combinations of: Orthogonal Projection Analysis (OPA), Alterating Least Squares (ALS), Simple-to-use Interactive Self-Modeling Mixture Analysis (SIMPLISMA) and VARIMAX. These are applied on the data to reduce tissue autofluorescence and resolve the spectral overlap in the emission spectra. The results show that spectral unmixing methods reduce the intensity caused by tissue autofluorescence by up to 78% and enhance image contrast by algorithmically reducing the overlap of the emission spectra.

***Keywords*—**breast carcinoma, hyperspectral imaging, m-FISH, spectral unmixing

I. Introduction

Breast carcinoma is the most common form of cancer in women. Almost 10% of malignancies in women are diagnosed as a *breast carcinoma*, which represents 22% of all cancer cases in women [1]. 5% to 10% of these *breast carcinoma* are genetically conditioned. The risk for women, whose mother or sister had a *breast carcinoma*, is twice that of women without a positive family anamnesis. This risk increases by a factor of between four and six if two family members developed cancer [2].Hence methods for a reliable diagnosis of *breast carcinoma* during routine checks are required.

Fluorescent in-situ hybridisation (FISH) is a technology that is used to stage *breast carcinoma*. FISH marks different cell components (e.g. nucleus, cytoplasm, proteins) as well as specific DNA positions or entire DNA sequences with fluorescently labeled DNA probes. Fluorochromes are substances that emit light, when excited by a specific wavelength. The emitted light has a longer wavelength than the excitation light. Fluorescence microscopy can be used to measure fluorescence and acquire images of FISH samples. It uses bandpass filters to measure the emission of the fluorochromes.

FISH samples marked with multiple fluorescently labeled DNA probes are termed multicolour-FISH (m-FISH). Analysis of (m-FISH) images is complicated by two problems: spectral overlap of the emission spectra and tissue autofluorescence. The spectral overlap is caused by the broad banded emission spectra of the fluorochromes, see fig. 1, which can not be resolved completely by emission filters or fluorochrome selection. Tissue autofluorescence originates from substances such as collagen or elastin. The substances show an unspecific broad-band fluorescence emission which overlaps the wanted signal and thus causes a decrease of the image quality.

Currently, pathologists use RGB colour images of m-FISH samples to make their diagnosis. However, the quality of the images makes diagnosis difficult and requires much experience for a reliable diagnosis. The quality of these images can be enhanced with spectral unmixing (SU) methods. A hyper-spectral imaging system, e.g. a tuneable filter mounted on a camera, measures the spectrum at each pixel in an image. The information content of these hyper-spectral images is higher than in standard colour images, enabling SU methods to unmix the overlapping emission spectra efficiently.

In this paper the following semi-supervised spectral unmixing methods are applied on hyper-spectral images of m-FISH samples: Principle Component Analysis (PCA), Orthogonal Projection Analysis (OPA), Alterating Least Squares (ALS), Simple-to-use Interactive Self-Modeling Mixture Analysis (SIMPLISMA) and VARIMAX. Results show a reduction of tissue autofluorescence by 78% and a contrast enhancement over standard RGB images.

II. Related Work

This section gives an brief overview of the related work in spectral imaging, detection of HER-2/neu status and spectral unmixing.

A. Spectral Imaging

Spectral imaging (SI) acquires spatially resolved images at different wavelengths and combines them into a three dimensional image cube. The two main approaches for the acquisition of hyper-spectral image data are *wavelength scanning* and spatial scanning.

Wavelength scanning methods take images for a certain wavelength range and both spatial axes at once. The spectral information is generated sequentially. *Wavelength scanning* systems

Martin De Biasio works as a R&D engineer in the field of spectral imaging for the Carinthian Tech Research AG, Villach, Austria, email: martin.debiasio@ctr.at

Raimund Leitner works as a program manager for spectral imaging systems at the Carinthian Tech Research AG, Villach, Austria, email: raimund.leitner@ctr.at

Sergey Verzakov works as a R&D engineer for Prime Vision, Delft, Netherlands, email: s.verzakov@primevision.com

Franz Wuertz works as a senior general pathologist at Klagenfurt hospital, Austria, email: Franz.Wuertz@lkh-klu.at

Pierre Elbischger is a professor of image processing and pattern recognition at Carinthian University of Applied Sciences, Klagenfurt, Austria, email: p.elbischger@cuas.at

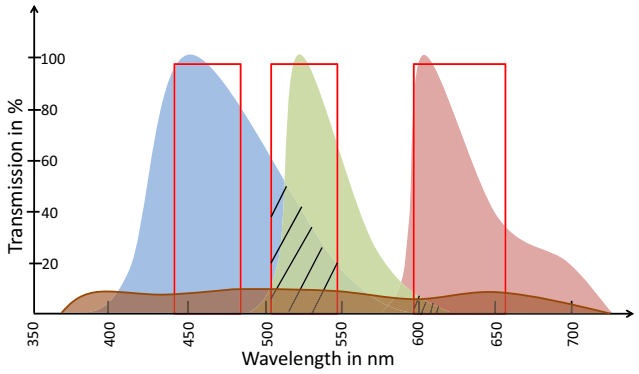


Fig. 1. Crosstalk in the emission spectra of three common fluorochromes DAPI (blue), FITC (green) and SpectrumOrange (red). The red rectangles represent an ideal filter curve of the emission filter. Even with ideal filters there is a crosstalk in the wavelength region between $500nm$ and $550nm$ where the fluorochromes DAPI and FITC heavily overlap. In the wavelength range from $600nm$ to $620nm$ the emission spectra of the fluorochromes FITC and SpectrumOrange overlap. The brown emission spectrum represents tissue autofluorescence. It is present over all wavelengths and reduces image quality.

can be built up using discrete optical filters or tunable filters. Filter based spectral imaging techniques use bandpass filters which are placed in front of a camera. By acquiring images, using different filters infront of the camera the hyper-spectral image data is acquired. The spectral resolution of the filter system is limited by the number of filters used. Liquid Crystal Tunable Filters (LCTF) or Acousto-Optical Tunable Filters (AOTF) are other *wavelength scanning* approach. Both are electronically tunable bandpass filters. The advantage of these components is that no mechanical motion during acquisition is involved.

Spatial scanning techniques such as imaging spectrographs are optical components that enable *spatial scanning* of measurement samples. Prism-grating-prism combinations disperse incident light of a single line of an object into its spectra and project it to a two dimensional sensor array. Spatial information is mapped to the x-axis, while spectral information is mapped on the y-axis. Hyper-spectral image data is generated by scanning the measurement sample linewise and combining the spectra of each line to a hyper-spectral image cube. When using imaging spectrographs, the spectral resolution is limited by the width of the entrance slit and the camera resolution. The advantage of this method is that each image pixel always contains the spectrum of the same object pixel. This is an important advantage over *wavelength scanning* approaches, where movements of the object during an acquisition cause the mixing of spectra of different objects in one image pixel.

B. Spectral Unmixing

A common way to resolve crosstalk in the emission spectra and reduce tissue autofluorescence of fluorescence measurements is spectral unmixing (SU). The method assumes that every pixel consists of a linear combination of overlapping emission spectra. There have been several implementations of spectral unmixing methods in the last years.

Munoz-Barrutia et. al. used Non-Negative Matrix Factorization (NMF) to blindly estimate spectral contributions in m-FISH spectral imaging data to correct the spectral overlap. Results showed, that this method outperforms approaches with prior knowledge about the spectra [7].

The HER-2/neu to CEP-17 ratio is an important factor for staging *breast carcinoma*. *Raimondo et. al.* developed an algorithm to determine HER-2/neu status for the classification of FISH images from *breast carcinomas*. The algorithm segments cell nuclei and FISH dots in two stages. For dot segmentation it uses a top-hat filtering stage followed by a template matching to separate real signals from noise[12]. For morphological analysis the authors use geometric rules to distinguish between holes within a nucleus and holes between neighboring nuclei. For overall case classification the algorithm calculates FISH signal ratio per cell nucleus and combines the results from multiple images from a slice [12].

III. Background

A brief introduction to *breast carcinoma* and the methods for classification and diagnostics of this life threatening disease are given here.

A. Epidemiology of breast Carcinoma

Breast carcinoma is caused by a malfunction in the cellular mechanisms which regulate growth [1]. Proto-oncogenes are normal genes that are responsible for the development and differentiation of cells. Mutations, such as point mutation, chromosome translocation and gene-amplification, can cause these proto-oncogenes to change their behaviour and become hyperactive and even non-physiological. HER-2/neu is a proto-oncogene belonging to the family of tyrosine kinasis receptors which has four subtypes HER-1, HER-2/neu, HER-3 and HER-4 [2]. These receptors are involved in the growth and the differentiation of cells. HER-2/neu is one of the few evidence-based features for the diagnosis of *breast carcinoma* [1]. Normal breast epithelial cells have two HER-2/neu gene copies and between 20,000 and 40,000 HER-2/neu receptors. In the early stages of 20% of *breast carcinomas* the HER-2/neu is overexpressed because of gene-amplification [3], [4], [5]. This increases the number of HER-2/neu receptors on the surface relative to normal breast epithelial cells [6]. The two steps for the detection and analysis of *breast carcinoma* are described in following paragraphs.

1) Fluorescent in-situ hybridisation (FISH): This is a common method to detect a HER-2/neu gene-amplification, deletions and translocations in tumour tissue [8]. The technique is based on the ability of single DNA strands to replicate themselves by recombining with a complementary base sequence to form a double strand (hybridisation). This double strand is then split into two single strands using temperature ranging from 70° to 95° [10]. This hybridisation is performed at the point where the probe and the DNA of the sample are complementary. A commercially available DNA probe, which is conjugated with a fluorochrome, is hybridised to a specific DNA sequence. Various companies provide special DNA probes for staining human tumour cells. These kits include

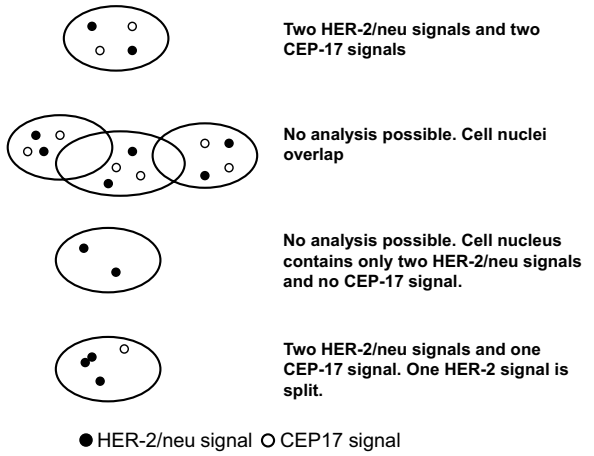


Fig. 2. Exemplary counting criterions for analysis of m-FISH samples recommended by the company Abbott [11].

the fluorochromes SpectrumOrange to hybridise HER-2/neu receptors, SpectrumGreen to hybridise chromosome No. 17 (CEP 17) and DAPI to counterstain the cell nucleus.

2) Analysis of breast carcinoma: The analysis of *breast carcinoma* samples is done with fluorescence microscopes, which are equipped with fluorescence filter combinations (e.g. DAPI, FITC and SpectrumOrange) to discriminate between different fluorochromes. The fluorescence emission peaks of HER-2/neu, CEP 17 and the cell nucleus are filtered with the excitation/emission filter combinations listed in Table I.

TABLE I
BANDPASS FILTER COMBINATIONS FOR FLUOROCHROMES

Cell component	Excitation filter CWL(FWHM)	Emission filter CWL (FWHM)
Cell nucleus	395nm (17nm)	461nm (15nm)
CEP 17	488nm (14nm)	523nm (19nm)
HER-2	562nm (24nm)	600nm (30nm)

For the analysis of the fluorescence samples the fluorescent signals of CEP 17 and HER-2/neu are of importance. The medical expert has to count the number of signals of 20 morphologically intact non-overlapping tumor cell nuclei with a clear fluorescence signal spot [11]. All cell nuclei need to have at least one CEP 17 and one HER-2/neu signal spot to be counted. Two signal spots of the same size must have a distance of at least the single fluorescence spot diameter to be counted as a single signal spot [10], compare fig. 2. Lymphocytes, granulocytes, macrophages, fibroblasts, epithelial cells, signals with low intensity, as well as tumor cells with no clear border or a high background signal need to be excluded from the analysis [10]. These are just exemplary criterion out of a huge range.

IV. METHODS

The evaluated spectral unmixing (SU) methods are discussed in this section.

A. Spectral Unmixing

The following notation of the variables for the explanation of the SU methods is used. $\mathbf{X}$ is a $M \times N$ matrix with M rows and N columns. Each row contains a spectrum and each column represents a wavelength. The rows of $\mathbf{X}$ are denoted by $\mathbf{x}_{\mathrm{i}}^T = (\mathbf{x}_{i,1}\mathbf{x}_{i,2}\ldots\mathbf{x}_{i,N})$ and represent distinct spectra. $\mathbf{Y}$ is a matrix containing K reference spectra $\mathbf{x}_{\mathrm{ref,k}}$. The reference spectra $\mathbf{x}_{\mathrm{ref,k}}$ are L2 normalised and are stored along the rows of the matrix $\mathbf{Y}$. K is defined by the user.

$$\mathbf{Y} = \begin{pmatrix} \mathbf{x}_{\mathrm{ref},1}^T \\ \mathbf{x}_{\mathrm{ref},2}^T \\ \vdots \\ \mathbf{x}_{\mathrm{ref,K}}^T \end{pmatrix} \tag{1}$$

The dispersion matrix

$$\mathbf{Y}_i = \begin{pmatrix} \mathbf{Y} \\ \mathbf{x}_i^T \end{pmatrix} \tag{2}$$

contains all normalised reference spectra of $\mathbf{Y}$ and the i-th spectrum $\mathbf{x}_i$ in the last row.

1) Orthogonal projection analysis (OPA): Orthogonal projection analysis (OPA) was proposed in [14]. The method iteratively determines reference spectra that optimise dissimilarity. The dissimilarity of a given $\mathbf{Y}_i$ is defined as:

$$d_i = \det(\mathbf{Y}_i\mathbf{Y}_i^T) = (\|x_{\mathrm{ref}}\|\|x_i\|\sin\alpha_i)^2, \quad i = 1,\ldots M. \tag{3}$$

$.^T$ is the transpose operator and the Euclidean norm is denoted by $||.||$. The reference spectrum and all spectra $\mathbf{x}_i$ in $\mathbf{Y}_i$ are normalised to unit length.
There are two ways to initialise the first reference spectrum $\mathbf{x}_{ref,1}$ of OPA. The first possibility is to pick a reference spectrum of the rows of $\mathbf{X}$ randomly. The second way is to calculate the mean spectrum $\overline{\mathbf{x}}$ of $\mathbf{X}$ by

$$\mathbf{x}_{\mathrm{ref},1} = \overline{\mathbf{x}}_{\mathrm{m}} = \frac{1}{M}\sum_{i=1}^{M}\mathbf{x}_i \tag{4}$$

and use it as the first reference spectrum $\mathbf{x}_{\mathrm{ref},1}$.
Thus, $\mathbf{Y}$ is initially $\mathbf{x}_{\mathrm{ref},1}$. d_i for all spectra $\mathbf{x}_i$ in $\mathbf{X}$ are calculated. The reference spectrum $\mathbf{x}_{\mathrm{ref}}$ and each spectrum $\mathbf{x}_i$ span up a parallelogram in the N-dimensional space. The area of the parallelogram is equal to the determinant of the dispersion matrix $\mathbf{Y}_i$. The value of the dissimilarity depends on the size of the area. The greater the value, the higher the dissimilarity.
The spectrum

$$\mathbf{x}_{\mathrm{ref},2} = \mathbf{x}_j, \quad j = \arg\max(d_i) \tag{5}$$

with the largest dissimilarity from the reference spectrum $\mathbf{x}_{\mathrm{ref},1}$ is selected and replaces the initial estimate for $\mathbf{x}_{\mathrm{ref},1}$

as a new reference spectrum into $\mathbf{Y}$. In the next iteration the determinant of $\mathbf{Y}_i$ is recalculated and the spectrum $\mathbf{x}_{\mathrm{ref},2}$ that yields the highest dissimilarity to $\mathbf{Y}_i$ is selected and added as the second reference spectrum to $\mathbf{Y}$.
In the next iteration, d_i of the dispersion matrix $\mathbf{Y}_i$ for $\mathbf{x}_i$ with respect to $\mathbf{x}_{\mathrm{ref},1}$ and $\mathbf{x}_{\mathrm{ref},2}$ is calculated. The procedure adds reference spectra to $\mathbf{Y}$ until the K predefined spectra are found or the relative mean squared error of the selected spectra as basis is smaller than a user defined threshold τ.

2) Simple-to-use interactive self-modeling mixture analysis (SIMPLISMA): Simple-to-use interactive self-modeling mixture analysis originates from [15]. Initially, when no spectrum has been selected $\mathbf{w}_i^T = (\mathbf{w}_{i,1}\mathbf{w}_{i,2}\ldots\mathbf{w}_{i,N})$ is initialised with ones. Normalisation of each spectrum by

$$\mathbf{z}_i = \frac{\mathbf{x}_i}{\sqrt{N(\mu_i^2 + (\sigma_i + \alpha)^2)}}. \tag{6}$$

causes that the length of the spectra containing a signal will be close to one. Spectra with low mean intensities i.e. noise spectra will be down weighted to zero. In equ. 6 σ_i and μ_i represent the standard deviation and mean of the i-th spectrum $\mathbf{x}_i$.
In contrast to OPA, SIMPLISMA does not use the mean spectrum $\overline{\mathbf{x}}$ to start the iterations but instead the 'most pure' spectrum. The purity $\mathbf{p}_i$ of the i-th spectrum $\mathbf{x}_i$ is defined by

$$\mathbf{p}_i = \mathbf{w}_i \frac{\sigma_i}{\mu_i + \alpha} \tag{7}$$

The offset value α is a small fraction of the largest spectral intensity of the whole dataset. The most pure spectrum is by definition the spectrum $\mathbf{x}_{\mathrm{ref},1}$ that maximises purity. When the first $\mathbf{p}_i$ has been selected

$$w_i = \det(\mathbf{Y}_i \mathbf{Y}_i^T) \tag{8}$$

is equal to the determinant of the dispersion matrix $\mathbf{Y}_i$. In case of SIMPLISMA $\mathbf{Y}_i$ consists of the selected pure spectra $\mathbf{z}_i$.
In principle SIMPLISMA has the same stopping criteria as OPA. The difference is that SIMPLISMA substitutes the dissimilarities by purities. Thus, SIMPLISMA differs from OPA only by specific scaling (normalisation) and the way how it starts iterations (initialisation).

3) VARIMAX: Another unmixing algorithm is VARIMAX [16], [17]. Instead of selecting the most dissimilar or pure spectra it extracts them, i.e. the reference spectra are linear combinations of all spectra from the dataset. The methods starts with principal component analysis (PCA) and retains the specified number of eigenvectors (loadings) $\mathbf{V}$. Loadings are also called abstract factors, because they are usually completely different from the real factors – pure spectra. The value $\mathbf{s}$ called 'simplicity' can be used to estimate how close this particular factor is to the pure spectrum. Simplicity $\mathbf{s}_i$ of the i-th spectrum is calculated by

$$\mathbf{s}_i = \frac{1}{N}\left(\sum_{i=1}^{N} \mathbf{x}_i'^4 \frac{\left(\sum_{i=1}^{N} \mathbf{x}_i'^2\right)^2}{N}\right). \tag{9}$$

VARIMAX iteratively performs planar rotations of the loading matrix $\mathbf{V}$ by

$$\mathbf{X'} = \mathbf{VX} \tag{10}$$

to maximise simplicity for the calculation of abstract loadings. The optimisation algorithm does not guarantee a global maximisation and thus has to be repeated several times.
There is a relative and absolute stopping criterion for VARIMAX. The absolute value is the difference $\mathbf{s}_i - \mathbf{s}_{i+1}$. The relative criterion is that the value which is optimised does not change at the next iteration more than some percentage of the previous value $(\mathbf{s}_i - \mathbf{s}_{i+1})/\mathbf{s}_i < \epsilon$.

4) Alterating least squares (ALS): Alterating least squares (ALS) is a purification algorithm. It is based on the assumption that spectral values and abundances are positive. Assumed a set of pure candidate spectra $\mathbf{Y}_0$ can be obtained e.g. by OPA. If they were definitely pure spectra then abundances could be found by solving the least squares problem

$$\mathbf{X} = \mathbf{Z}_1\ \mathbf{Y}_0. \tag{11}$$

$\mathbf{Z}_1$ is the pure concentration of a component and is calculated by

$$\mathbf{Z}_1 = \mathbf{X}\mathbf{Y}_0^T(\mathbf{Y}_0\mathbf{Y}_0^T)^{-1}. \tag{12}$$

However, in practice one has to deal with imperfect candidate spectra. Simple factors are orthogonal and usually contain positive and negative values, leading to the situation that some abundances obtained by the least squares solution will be almost for sure negative. In ALS they are clipped to zero $\mathbf{Z}_{1,\mathrm{c}}$ and used again to recompute spectra candidates with Eqn. (12). $\mathbf{X}$ can also contain negative spectral values. With

$$\mathbf{X} = \mathbf{Z}_{1,\mathrm{c}}\mathbf{Y}_1 \tag{13}$$

and

$$\mathbf{Y}_1 = \mathbf{Z}_{1,\mathrm{c}}^T\mathbf{X}(\mathbf{Z}_{1,\mathrm{c}}^T\mathbf{Z}_{1,\mathrm{c}})^{-1}. \tag{14}$$

they are clipped to zero resulting in $\mathbf{Y}_{1,c}$. The algorithm is iterated until the convergence criterion

$$d\mathbf{Y}_1 = \mathbf{Y}_{1,\mathrm{c}} - \mathbf{Y}_0 < \epsilon \tag{15}$$

is met. If the algorithm converged or the number of defined iterations is reached then $\mathbf{Y}$ is set to $\mathbf{Y}_{1,\mathrm{c}}$ and $\mathbf{Z}$ is set to $\mathbf{Z}_{1,\mathrm{c}}$. Otherwise $\mathbf{Y}_0$ is set to $\mathbf{Y}_{1,\mathrm{c}}$ and iterations go on.

V. Data Acquisition

For the evaluation of the SU methods described in section IV-A, hyper-spectral images of m-FISH samples were acquired. An Axio Imager A1m (Carl Zeiss, Germany) fluorescence microscope, in combination with a triple bandpass filter combination no. 40 DAPI/FITC/Cy3™ (Chroma, USA), was used to acquire m-FISH data sets. A liquid crystal tuneable filter (LCTF) (Varispec™, CRi, MA, USA) was used to acquire images from $400\mathrm{nm}$ to $720\mathrm{nm}$ with a step width of $5\mathrm{nm}$.
The samples were excitated with a metal halide lamp X-Cite 120PC (EXFO®, Canada). A high sensitivity 14 bit EMCCD camera iXon (Andor, Ireland) with a pixel resolution of 1004×1002 pixels was used to acquire the images.
A spatial two-point calibration for offset and sensitivity of each pixel of the camera was done with a dark image (closed aperture) and a white reference. This calibration model corrects the spatial and spectral characteristics of the measurement setup with a linear model.

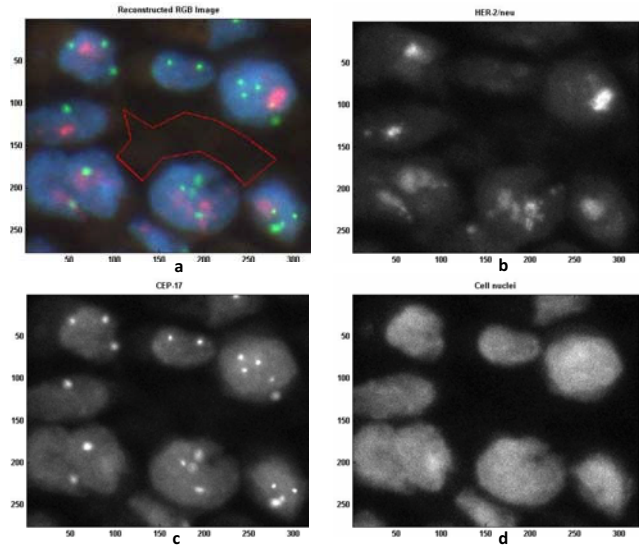


Fig. 3. Reconstructed colour image from the wavelengths $470\mathrm{nm}$, $530\mathrm{nm}$ and $625\mathrm{nm}$ from hyper-spectral image data of a *breast carcinoma* sample marked with m-FISH probes (a). The sample shows HER-2/neu gene amplifications in all cell nuclei. (b), (c) and (d) show the fluorescent spots of HER-2/neu genes, CEP 17 genes and the cell nuclei. The red polygon in (a) represents the user defined ROI for the analyse of tissue autofluorescence reduction.

Data acquisition was carried out in two steps. First, regions with a sufficient number of unconnected cell nuclei were identified with a magnification of 100 fold. Second, images of these regions were acquired at a magnification of 1,000. Focal distance varies with wavelength, hence the focus was adjusted for the three emission wavelength ranges ($400\mathrm{nm}$-$500\mathrm{nm}$, $500\mathrm{nm}$-$600\mathrm{nm}$, $600\mathrm{nm}$-$720\mathrm{nm}$). Images were acquired for each wavelength region separately.
For data analysis eight m-FISH data sets with 64 single images each were acquired. Fig. 3a shows a reconstructed pseudo colour image from the wavelengths $470\mathrm{nm}$, $530\mathrm{nm}$ and $625\mathrm{nm}$ from hyper-spectral imaging data of a *breast carcinoma* sample hybridised with m-FISH probes. The image can be compared with a standard colour image which is routinely used by pathologists for the detection of *breast carcinoma*. Fig. 3b and 3c show fluorescence spots of HER-2/neu genes and CEP 17 genes. The faint regions that surround the bright fluorescent spots correspond to the overlapping fluorescence signal of DAPI marking the cell nuclei, see Fig. 3d. This effect is caused by the overlapping emission spectrum of the fluorescent probes for the cell nuclei, with the ones of HER-2/neu genes and CEP 17 genes.

VI. Results

The SU algorithms introduced in section IV-A were used to reduce tissue autofluorescence and enhance image contrast. In practice combining the single SU methods gives better results than using the methods separately. The following combinations were used to unmix hyper-spectral m-FISH images:

- OPA/ALS
- SIMPLISMA/ALS
- PCA/VARIMAX/ALS

In case of the combination OPA/ALS, OPA is applied first. The results of OPA are then the starting point for ALS. All of these algorithms are semi-supervised; i.e. whereby the user has to set the maximum number of components to be unmixed.

A. Reduction of tissue autofluorescence

Tissue autofluorescence degrades image quality. It has a nonspecific emission spectrum that adds to the intensity in every channel. It falsifies the wanted signal information and is therefore disturbing.
For a given region of interest (ROI) $\mathcal{N}$ the mean tissue autofluorescence is calculated by

$$\overline{\mathrm{AF}}_{\mathcal{N}}(\lambda_{\mathrm{N}}) = \frac{1}{|\mathcal{N}|} \sum_{\mathrm{x} \in \mathcal{N}} \mathrm{AF}(\mathrm{x_k}, \mathrm{y_k}, \lambda_{\mathrm{N}}), \qquad (\mathrm{x}, \mathrm{y}) \epsilon \mathcal{N} \quad (16)$$

The equation takes every pixel x_{k}, y_{k} in a given ROI $\mathcal{N}$ and calculates the mean and standard deviation of intensity for each of the channels λ_{n} of the image.
Table II lists the mean values of tissue autofluorescence for each channel achieved with the different SU methods. The results of the three SU combinations are compared with standard RGB images. Standard RGB images were generated by summing up all images in the specific wavelength range and dividing it by the number of images. The blue channel consists of wavelengths from $450\mathrm{nm}$ to $480\mathrm{nm}$, the green channel contains wavelengths between $515\mathrm{nm}$ and $555\mathrm{nm}$ and the red channel ranges between $595\mathrm{nm}$ and $660\mathrm{nm}$.
The combination PCA/VARIMAX/ALS gave the best results. Compared to the standard RGB image a reduction of tissue autofluorescence of 80% for HER-2/neu spots, 64% for CEP 17 spots and 90% for cell nuclei spots was achieved.

B. Enhancement of image contrast

Overlapping emission spectra of fluorochromes can not be completely resolved with the used emission filter. Due to their

TABLE II
MEAN VALUES OF SIX IMAGES FOR EACH COLOUR CHANNEL OF USER DEFINED ROIS.

	Tissue autofluorescence for used camera channel		
Dataset	**Red Mean (Std)**	**Green Mean (Std)**	**Blue Mean (Std)**
Standard RGB Image	0.15 (0.01)	0.14 (0.01)	0.1 (0.03)
OPA/ALS	0.08 (0.04)	0.11 (0.04)	0.03 (0.02)
SIMPLISMA/ALS	0.11 (0.02)	0.11 (0.02)	0.02 (0.01)
PCA/VARIMAX/ALS	0.03 (0.01)	0.05 (0.01)	0.01 (0.01)

spectral overlap of the fluorochromes each pixel in an hyperspectral image cube can be interpreted as being either the cell nuclei hybridised with DAPI; CEP 17 genes hybridised with FITC; HER-2/neu genes hybridised with SpectrumOrange or tissue autofluorescence.

Fig. 4a shows a 3D plot of a standard RGB image routinely used by pathologists. It is obvious that there is a large overlap of the four point clouds (cell nuclei, CEP 17, HER-2/neu, background) of the fluorescence spots. Fig 4b shows a 3D plot of an unmixed image achieved with the combination PCA/VARIMAX/ALS. In this case the point clouds of cell nuclei, CEP 17, HER-2/neu and background pixels are better separated. There is just a small percentage of overlapping pixels.

For the enhancement of image contrast, a standard RGB image was compared with the results of the SU methods combinations. A fixed intensity value threshold of greater 0.5 for cell nuclei, CEP 17 and HER-2/neu pixels was used. For the background the threshold was set to an intensity value lower 0.3. For the analysis all pixels in the image, which were ambiguously assigned as cell nuclei, CEP 17 or HER-2/neu pixels were counted.

TABLE III
PIXEL MISCLASSIFICATION FOR THREE M-FISH DATASETS.

Unmixing result	**DS I %**	**DS II %**	**DS III %**	**Overall mean (std)**
Standard RGB Image	44.1	6.94	14.33	21.79 (19.67)
OPA/ALS	0.04	5.3	2.37	2.57 (2.64)
SIMPLISMA/ALS	0.18	4.69	2.04	2.3 (2.27)
PCA/VARIMAX/ALS	0.039	1.96	1.23	1.08 (0.97)

Table III lists the results of the analysis of three m-FISH images. In a standard RGB image in mean 21.79% of all pixels can not be assigned uniquely to one class. When using SU methods the percentage of ambiguously assigned pixels is reduced. The combinations OPA/ALS and SIMPLISMA/ALS reduce the number to 2.57%. The similar results for both combinations can be explained by the small differences between the two SU methods SIMPLISMA and OPA. The best result with 1.08% is achieved with the combination PCA/VARIMAX/ALS.

The high variances in the three analysed images is caused by the differences of the fluorescent signal intensity of the fluorescent probes. Some of the fluorescently labeled m-FISH images were already bleached before the experiment.

VII. CONCLUSION

Spectral unmixing methods have been applied to hyperspectral m-FISH data. The objective of the analysis was to reduce tissue autofluorescence and to enhance image contrast. The unmixing results were compared with standard RGB images using both of these criteria. It was shown that the combination SIMPLISMA/ALS reduced tissue autofluorescence by 80% for HER-2/neu fluorescence spots, 64% for CEP 17 fluorescence spots and 90% for cell nuclei fluorescence spots. For the enhancement of image contrast the percentage of pixels which were ambiguously assigned to be rather HER-2/neu, CEP 17 or cell nuclei pixels were analysed. In a standard RGB image 21.8% of all pixels were ambiguously assigned. This value was reduced to 1.08% with the combination PCA/VARIMAX/ALS.

This overall enhancement of image quality simplifies automated fluorescence spot counting algorithms and thus enables a more reliable detection of *breast carcinoma* using m-FISH stained histological preparations.

ACKNOWLEDGEMENT

The authors would like to thank Mrs. Christine Perkonigg from Klagenfurt Hospital Department of Pathology for providing the FISH samples.

REFERENCES

[1] N. Harbeck, W. Eiermann, J. Engel, I. Funke, A. Lebeau, W. Parmanetter and M. Untch *Prognosefaktoren beim primären Mammakarzinom*, Tumorzentrum Muenchen: Manual Mammakarzinome: Empfehlungen zur Diagnostik, Therapie und Nachsorge, 8. Aufl., Zuckerschwedt, Muenchen, Bern, Wien, New York, pages 39-43, 2001

[2] R. Lupu, M. Cardillo, L. Harris, M. Hijazi and K. Rosenberg *Interaction between erbB-receptors and heregulin in breast cancer tumor progression and drug resistance*, Sem Cancer Biol, Vol 6, pages 135-145, 1995

[3] D.M. Barnes, J. Bartkova, R.S. Camplejohn, W.J. Gullick, P.J. Smith P.J and R.R. Millis *Overexpression of the c-erbB-2 oncogene: Why does this occur more frequently in ductal carcinoma in situ than in invasive mammary carcinoma and is this of prognostic significance?*, Eur J Cancer, Vol 28, pages 644-648, 1992

[4] E. Liu, A. Thor, M. He, M. Barcos, B.M. Ljung and C. Benz *The HER2 (c-erbB-2) oncogene is frequently amplified in in situ carcinomas of the breast*, Oncogene, Vol 7, pages 1027-1032, 1992

[5] N.E. Hynes and D.F. Stern *The biology of erbB-2/neu/HER-2 and its role in cancer*, Biochim Biophys Acta, Vol 198, pages 165-184, 1994

[6] G.D. Lewis, I. Figari, B. Fendly, W.L. Wong, P. Carter, C. Gorman and H.M. Shepard *Differential responses of human tumor cell lines to anti-p185HER2 monoclonal antibodies*, Cancer Immunol Immunother, Vol 37, pages 255-263, 1993

[7] A. Munoz-Barrutia, J. Garcia-Munoz, B. Ucar, I. Fernandez-Garcia and C. Ortiz-de-Solorzano *Semi-supervised linear spectral unmixing using a hierarchical Bayesian model for hyperspectral imagery* , IRIT/ENSEEIHT/TeSA, 2 rue Camichel, BP 7122, 31071 Toulouse cedex 7, France, 2007

[8] E. Hyytinen, T. Visakorpi, A. Kallioniemi, O.P. Kallioniemi and J. Isola *Improved technique for analysis of formalin-fixed, paraffin-embedded tumors by fluorescence in situ hybridization*, Cytometry, Vol 16, pages 93-99, 1994

[9] J. Chang-Claude, H. Becher, U. Hamann and T. Schroeder-Kurtz *Risikoabschätzung fuer das familiaere Auftreten von Brustkrebs*, Zentralblatt fuer Gynaekologie 117, pages 432-434, 1995

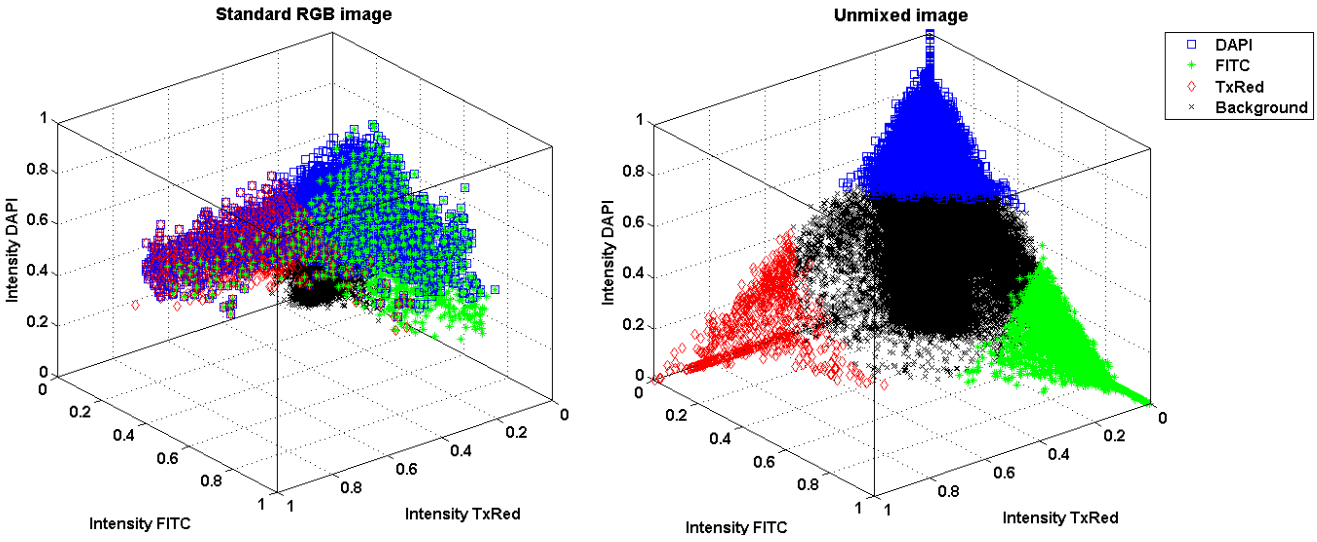


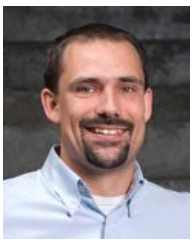
Fig. 4. 3D plot of the SU result achieved with the SU method combination PCA/VARIMAX/ALS of a hyper-spectral m-FISH *breast carcinoma* tissue sample. In both 3D plots blue squares represent Cell nuclei spots, green stars represent CEP 17 genes spots, red diamonds represent HER-2/neu gene spots and black x represent the background. There is a heavy overlap of fluorescent spots in the standard RGB image, see (a). (b) shows that the overlap is reduced by SU methods. The four point clouds in the 3D plot are in the corners of the plot and only a few points overlap.

[10] S. Herzog *Evaluation der HER-2/neu-Genamplifikation an Tumorabklatschpräparaten von Mammakarzinomen mittels FISH und deren Wertigkeit im Vergleich zur immunhistochemischen membranassoziierten Proteinueberexpression*, Universitäts-Frauenklinik mit Poliklinik Tuebingen, Allgemeine Geburtshilfe und Frauenheilkunde, Germany, pages 22-25, 2005
[11] *PathVysion* HER-2/neu DNA Probe Kit Package Insert, Vysis Inc.
[12] F. Raimondo, MA Gavrielides, G. Karayannopoulou, K. Lyroudia, I. Pitas and I. Kostopoulos *Automated evaluation of Her-2/neu status in breast tissue from fluorescent in situ hybridization images*, IEEE Trans Image Processing 14(9), pages 1288-99, Sep. 2005
[13] N. Dobigeon, J. Tourneret and Ch. Chang *Blind Spectral Unmixing of m-FISH Images by Non-negative Matrix Factorization*, Engineering in Medicine and Biology Society, 2007, EMBS 2007, 29th Annual International Conference of the IEEE, Page(s):6247 - 6250, Volume, Issue, 22-26 Aug. 2007
[14] F. C. Snachez, J. Toft, B. van den Bogaert and D. L. Massart *Orthogonal Projection Approach Applied to Peak Purity Asessment*, Anla. Chem. 68, 1,pages 79-85, 1996
[15] W. Windig and J. Guilment, *Interactive Self-Modeling Mixture Analysis*, Anla. Chem. 63, pages 1425-1432, 1991
[16] F. H. Kaiser *The varimax Criterion for Analytic Rotation in Factor Analysis*, Psyhometrika 23, pages 187-200, 1958
[17] M. Forina, C. Armanio, S. Lanteri and R. Leardi *Methods of VARIMAX rotations ini factor analysis with applications in clinical and food chemistry*, Journal of Chemometrics, Vol. 2, pages 115-125, 1998

Martin De Biasio received his Dipl.Ing. (FH) degree (with honors) in Medical Information Technology from Carinthia University of Applied Sciences (CUAS) Klagenfurt, Austria, in 2007. Currently he is working fulltime for the Carinthian Tech Research AG in the field of medical and industrial hyper-spectral image processing. He is also doing his master of science degree in Health Care IT at CUAS. His research interests include multicolour fluorescence microscopy and hyper-spectral imaging.

Franz G. Wuertz was borne in Vienna, Austria, 1957. He studied medicine at the medical university of Vienna. From 1981 to 1982 he was employed at the viennese hospital Hietzing. Since 1983 he works at the department of pathology at the general hospital of Klagenfurt, Austria, as a senior general pathologist with emphasis on Immunofluorescence-tests, FISH, Hematopathology and Neurooncopathology.

Raimund Leitner received his MSc degree in Telematics from Graz University of Technology in 2001. The two main fields of his Telematics study were machine vision and hardware/software development. In 2001 he started as researcher for industrial machine vision and spectral imaging at Carinthian Tech Research CTR AG in Villach, Austria. In his doctorate started in 2002 he worked part-time on unsupervised learning for object recognition of complex 3D objects which was finished in 2007. During his work at CTR he has led several industrial and research projects in the field of machine vision and spectral imaging. Currently he is program manager for spectral imaging and since 2008, when CTR was granted the K1 centre for advanced sensor technologies, also area manager for CTR's K1 area optical system technology. His research interests are machine vision, pattern recognition and spectral imaging.



Pierre J. Elbischger has been a member of IEEE since 2006. He received his MSc degree in Telematics from Graz University of Technology (GUT) in 2001. In his undergraduate studies he focused on discrete signal processing algorithms and digital signal processor architectures. Subsequently, his doctoral thesis was performed in the field of automated microscopy at the Institute of Computer Graphics and Vision at GUT, Austria, where he received his PhD in 2004. From 2004 to 2007 he was a post-doctoral fellow at School of Medical Information Technology, Carinthia University of Applied Sciences (CUAS), Austria, working in the field of ultrasound strain imaging. Currently, he is a professor of medical image processing and pattern recognition at CUAS, Austria.

Sergey Verzakov received his MSc degree in physics from Moscow Institute of Physics and Technology in 1996. From 1996 to 1999 he was a junior researcher at the institute of spectroscopy of Russian Academy of Science in a field of computer simulation of phase transitions. In 1999 he started collaboration with Chemometrics Group of Institute of Chemo- and Biosensors, Muenster, Germany. From 2003 to 2007 he was a PhD student at Delft University of Technology, The Netherlands. During this period his research topic was the analysis of hyperspectral images. Currently he is an employee of the pattern recognition company Prime Vision, The Netherlands and a guest researcher at Delft University of Technology.