Lecture VI
Dependent component analysis (DCA)

Ivica Kopriva

Ruđer Bošković Institute

e-mail: ikopriva@irb.hr ikopriva@gmail.com
Web: http://www.lair.irb.hr/ikopriva/

Course outline

- Motivation with illustration of applications (lecture I)
- Mathematical preliminaries with principal component analysis (PCA)? (lecture II)
- Independent component analysis (ICA) for linear static problems: information-theoretic approaches (lecture III)
- ICA for linear static problems: algebraic approaches (lecture IV)
- ICA for linear static problems with noise (lecture V)
- Dependent component analysis (DCA) (lecture VI)

Course outline

- Underdetermined blind source separation (BSS) and sparse component analysis (SCA) (lecture VII/VIII)
- Nonnegative matrix factorization (NMF) for determined and underdetermined BSS problems (lecture VIII/IX)
- BSS from linear convolutive (dynamic) mixtures (lecture X/XI)
- Nonlinear BSS (lecture XI/XII)
- Tensor factorization (TF): BSS of multidimensional sources and feature extraction (lecture XIII/XIV)

Outline

- Blind source separation (BSS) and independent component analysis (ICA)
- Preprocessing transforms for BSS – enhancing statistical independence
- Examples
  - separation of the images of human faces (seminar)
  - "model-free" blind image deconvolution (seminar ?)
  - "model-free" blind speech deconvolution (seminar ?)
  - blind segmentation of multispectral image of the tumor (seminar)
  - NIR and NMR spectroscopy

Seminar problems

1. Blind separation of two uniformly distributed signals with maximum likelihood (ML) and AMUSE/SOBI independent component analysis (ICA) algorithm.
   Blind separation of two speech signals with ML and AMUSE/SOBI ICA algorithm. **Theory, MATLAB demonstration and comments of the results.**

2. Blind decomposition/segmentation of multispectral (RGB) image using ICA, dependent component analysis (DCA) and nonnegative matrix factorization (NMF) algorithms. **Theory, MATLAB demonstration and comments of the results.**

3. Blind separation of acoustic (speech) signals from convolutive dynamic mixture. **Theory, MATLAB demonstration and comments of the results.**

Seminar problems

4. Blind separation of images of human faces using ICA and DCA algorithms (innovation transform and ICA, wavelet packets and ICA) **Theory, MATLAB demonstration and comments of the results.**

5. Blind decomposition of multispectral (RGB) image using sparse component analysis (SCA): clustering + L_p norm ($0<p\leq 1$) minimization . **Theory, MATLAB demonstration and comments of the results.**

6. Blind separation of four sinusoidal signals from two static mixtures (a computer generated example) using sparse component analysis (SCA): clustering + L_p norm ($0<p\leq 1$) minimization in frequency (Fourier) domain. **Theory, MATLAB demonstration and comments of the results.**

Seminar problems

7. Blind separation of three acoustic signals from two static mixtures (a computer generated example) using sparse component analysis (SCA): clustering + L_p norm ($0<p\leq1$) minimization in time-frequency (short-time Fourier) domain. **Theory, MATLAB demonstration and comments of the results.**

8. Blind extraction of five pure components from mass spectra of two static mixtures of chemical compounds using sparse component analysis (SCA): clustering a set of single component points + L_p norm ($0<p\leq1$) minimization in m/z domain. **Theory, MATLAB demonstration and comments of the results.**

9. Feature extraction from protein (mass) spectra by tensor factorization of disease and control samples in joint bases. Prediction of prostate/ovarian cancer. **Theory, MATLAB demonstration and comments of the results.**

BSS and ICA

A theory for multichannel blind signal recovery requiring minimum of *a priori* information.

Problem:

$$\mathbf{x}=\mathbf{A}\mathbf{s}$$

Goal: find **s** based on **x** only (**A** is unknown).

By means of independent component analysis (ICA) algorithms **W** can be found such that:

$$\mathbf{y}\cong\mathbf{s}=\mathbf{W}\mathbf{x} \rightarrow \mathbf{y} \cong \mathbf{P}\Lambda\mathbf{s}$$

[1]A. Hyvarinen, J. Karhunen, E. Oja, "Independent Component Analysis," John Wiley, 2001.
[2]A. Cichocki, S. Amari, "Adaptive Blind Signal and Image Processing," John Wiley, 2002.

When does ICA work !?

- source signals $s_i(t)$ must be statistically independent.

$$\mathrm{p}(\mathbf{s}) = \prod_{i=}^{N} p_i(s_i)$$

- source signals $s_i(t)$, except one, must be non-Gaussian.

$$C_n(s_i) \neq 0 \qquad n > 2$$

- mixing matrix A must be nonsingular.

$$\mathbf{W} \cong \mathbf{A}^{-1}$$

Understanding origin i.e. interpretation of the mixing matrix is sometimes of crucial importance.

Preprocessing transforms for BSS – enhancing statistical independence

- We want to find a linear operator T with the property that $T(s_m)$ and $T(s_n)$ are more independent than s_m and s_{n} $\forall m, n$.

- Then, $\mathbf{W}\cong\mathbf{A}^{-1}$ is learnt by applying ICA on $T(\mathbf{x})=\mathbf{A}T(\mathbf{s})$.

- The challenge is how to find appropriate linear operator T ?

Preprocessing transforms for BSS – enhancing statistical independence

- Sub-band decomposition ICA (SDICA): wideband source signals are dependent, but there exists some sub-bands where they are less dependent.[2,3-6]

- Innovations-based approach.[7]

- Adaptive pre-filtering of measured signals.[8,9]

[3] A. Cichocki, P. Georgiev, Blind source separation algorithms with matrix xonstraints, IEICE Trans. Fund. Electron. Commun. Comput. Sci. E86-A (2003) 522-531.
[4] T. Tanaka, A. Cichocki, Subband decomposition independent component analysis and new performance criteria, Proc. ICASSP, 2004
[5] I. Kopriva, D. Sersic, Wavelet packets approach to blind separation of statistically dependent sources, Neurocomputing **71**, 1642-1655 (2008).
[6] I. Kopriva, D. Sersic, Robust blind separation of statistically dependent sources using dual tree wavelets, ICIP 2007.
[7] A. Hyvarinnen, Independent component analysis for time-dependent stochastic processes, ICANN'98, Skvode, Sweden, 1998.
[8] K. Zhang, L.W. Chan, An adaptive method for subband decomposition ICA, Neural Computation 18 (2006) 191-223.
[9] K. Zhang, L.W. Chan, Enhancement of source independence for blind source separation, LNCS 3889 (2006) 731-738.

Why/when source signals are statistically dependent?

• Slow varying part of sources signals s_m and s_n is what makes them statistically dependent up to some extent. That is also known as trend. In term of frequency spectrum of s_m and s_n it is the similarity between the low-frequency parts of the spectra that makes them statistically dependent.

•Thus, removing of low-frequency part of the spectra will make source signals s_m and s_n less dependent. That is implemented by some type of the high-pass filter.

• High-pass filters can be either fixed (predefined) or data adaptive. They are linear operators and can be seen as approximations of differential operator that is use for trend removal:

$$\frac{d}{dt}\left(\mathbf{x}(t)=\mathbf{A}\mathbf{s}(t)\right):\frac{d\mathbf{x}(t)}{dt}=\mathbf{A}\frac{d\mathbf{s}(t)}{dt}$$

Increasing statistical independence

- First order difference approximation of differential operator yields:

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{A}\frac{d\mathbf{s}(t)}{dt} \rightarrow \mathbf{x}(kT) - \mathbf{x}((k-1)T) = \mathbf{A}\mathbf{s}(kT) - \mathbf{s}((k-1)T)$$
$$\rightarrow \tilde{\mathbf{x}}(kT) = \mathbf{A}\tilde{\mathbf{s}}(kT)$$

That amounts to preprocess measured data **x**(t) by first order high pass (HP) filter with coefficients [1 –1]. It is very simple to implement. However, more sophisticated approximation of derivative operator can be used that will yield different kinds of HP filters. They are characterized by impulse response of appropriate order:

$$\tilde{\mathbf{x}}(kT) = \sum_{m=0}^{M} h(mT)\mathbf{x}(kT - mT)$$
$$= \mathbf{A}\sum_{m=0}^{M} h(mT)\mathbf{s}(kT - mT) = \mathbf{A}\tilde{\mathbf{s}}(kT)$$

Increasing statistical independence

- After HP filtering. ICA is applied to:

$$\tilde{\mathbf{x}}(kT) = \mathbf{A}\tilde{\mathbf{s}}(kT)$$

That yields more accurate estimation of **A**. Sources are obtained as: $\mathbf{s}=\mathbf{A}^{-1}\mathbf{x}$.

•HP filtering yields signals that are *less dependent* and contain fast varying part of the signals. Thus, such signals are *more non-Gaussian* (super-Gaussian) than original signals. That is exactly what is required by ICA.

•However, HP filters considered so far are predefined (fixed).

Innovations (data adaptive)-based approach[7]

•Innovations (prediction errors) are also *more independent* from each other and *more non-Gaussian* than original processes → *essentially important* for the success of the ICA algorithms. They can be seen as data adaptive HP filters.

$$\begin{aligned}
\tilde{\mathbf{x}}(k) &= \mathbf{x}(k) - E\left[\mathbf{x}(k) \middle| \mathbf{x}(k-1), \mathbf{x}(k-2), \ldots\right] \\
&= \mathbf{A}\mathbf{s}(k) - E\left[\mathbf{A}\mathbf{s}(k) \middle| \mathbf{A}\mathbf{s}(k-1), \mathbf{A}\mathbf{s}(k-2), \ldots\right] \\
&= \mathbf{A}\left[\mathbf{s}(k) - E\left[\mathbf{s}(k) \middle| \mathbf{s}(k-1), \mathbf{s}(k-2), \ldots\right]\right] = \mathbf{A}\tilde{\mathbf{s}}(k)
\end{aligned}$$

•The approach is computationally efficient. Linear time invariant prediction error filter is efficiently estimated by means of Levinson algorithm (MATLAB command `lpc`).

Innovations (data adaptive)-based approach[7]

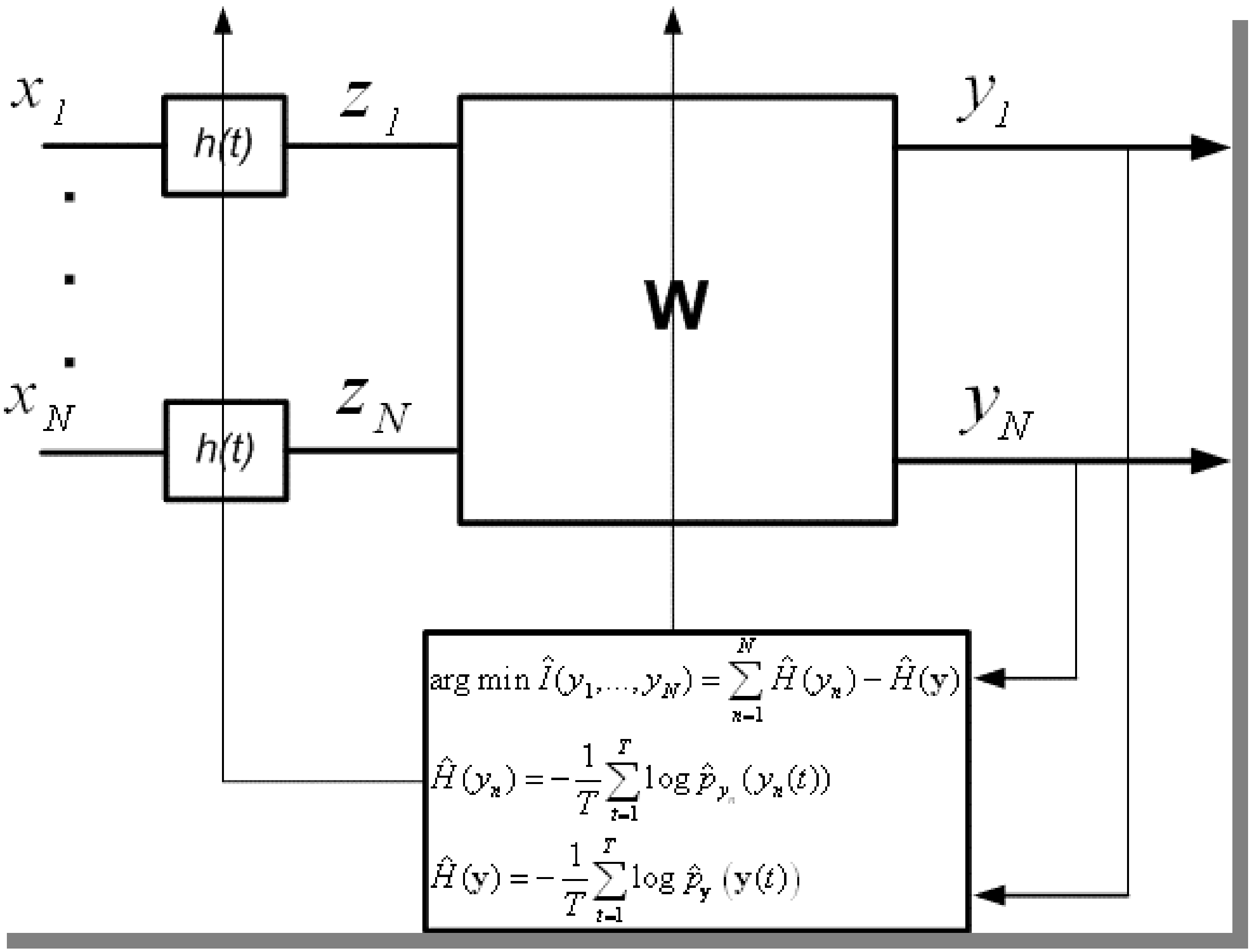
•The innovation approach requires source signals to have temporal (spatial) structure i.e. to be predictable. Thus, such approach makes sense, for statistically dependent sources, since dependence is caused by slow varying part of the signals.

• Prediction error filters are learned for each measured signal x_n $n=1,\ldots,N$ separately. However, to preserve phase lock between innovations of x_n the filter has to be applied to all x_n $n=1,\ldots,N$.

If $\mathbf{h}_n$ are impulse response of individual innovation filters the innovation filter used to get innovation of $\mathbf{x}$ is obtained as:

$$\mathbf{h} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{h}_n$$

Adaptive pre-filtering approach[8,9]



Adaptive pre-filtering approach[8,9]

• Measured signals are preprocessed by an adaptive filter with adaptation based on the minimum of mutual information (MI) among the restored sources.

•In the first version of the method, [9], prefilter and static ICA are implemented simultaneously. It has been found found in [5] that learning is not stable.

•In the improved version of the method, [9], prefilter is adapted on the subset of source signals, that are assumed to be available. ICA is then applied on filtered observed data. Assumption about availability of the subset of source signals is somewhat unrealistic in the BSS scenario.

Sub-band decomposition approach[3,4]

• In SDICA approach the operator *T* represents prefilter applied to all observed signals. In [3, 4], a fixed filter bank is used for signal filtering. By assumption the wideband source signals are dependent, but some of their subcomponents are less dependent: $\mathbf{s}(t)=\sum_{l=1}^{L}\mathbf{s}_l(t)$.

•The challenge is how to find a subband index $1\leq k\leq L$, such that $\mathbf{s}_k$ contains least dependent subcomponents? As a criteria for sub-band localization is vicinity of the global separation matrix $\mathbf{Q}=\mathbf{W}_m\mathbf{W}_n^{-1}$ to the general permutation matrix $\mathbf{P}=\mathbf{D}\Lambda$:

$$P_{err}=\frac{1}{N(N-1)}\sum_{i=1}^{N}\left\{\left(\sum_{j=1}^{N}\frac{|q_{ij}|}{\max_k|q_{ik}|}-1\right)+\left(\sum_{j=1}^{N}\frac{|q_{ji}|}{\max_k|q_{ki}|}-1\right)\right\}$$

Hence, <u>at best</u> two sub-bands can be localized.

Sub-band decomposition approach[5,6]

- In [5][6] it has been respectively proposed the use of real and complex discrete wavelet transform to implement multiscale sub-band decomposition.

- Use of wavelet transform enables implementation of multi-resolution analysis (MRA) decomposing signal into coarse (slow varying) and details (fast varying) parts, whereas a wavelet function to which signal is projected can be selected to match the type of signal.

- The continuous wavelet transform (CWT) of signal *s*(t) is:

$$CWT(a,\tau) = \frac{1}{\sqrt{a}} \int s(t)\psi\left(\frac{t-\tau}{a}\right)dt$$

Y. T. Chan, Wavelet Basics, KAP, 1995.
G. Strang, T. Nguyen, Wavelets and Filter Banks, Wellesley-Cambridge Press, 1996.
S. Mallat, A Wavelet Tour in Signal Processing, Academic Press, 1998.

Sub-band decomposition approach[5,6]

ψ(t) is the basic (or mother) wavelet and ψ((t-τ)/a)/sqrt(a) the wavelet basis functions (sometimes called baby wavelets). By change of variables at'=t we also obtain CWT:

$$CWT(a,\tau) = \sqrt{a}\int s(at')\psi\left(t' - \frac{\tau}{a}\right)dt'$$

That shows equivalence between scaling ψ(t) and scaling *s*(t) to obtain CWT. The CWT is implemented in MATLAB by the command `cwt`. For its use, as well as for other kinds of wavelet transforms, a wavelet toolbox is required. CWT can be seen in four different ways:

•(i) it computes the inner product (projection) or the cross-correlation of *s*(t) with ψ(t/a)/sqrt(a) at shift τ/a. It computes the component of *s*(t) that is common to ψ(t/a)/sqrt(a).

Sub-band decomposition approach[5,6]

•(ii) it is the output of a bandpass filter of impulse response $\psi(-t/a)/\text{sqrt}(a)$ with input *s*(t), at the time instant τ/a.

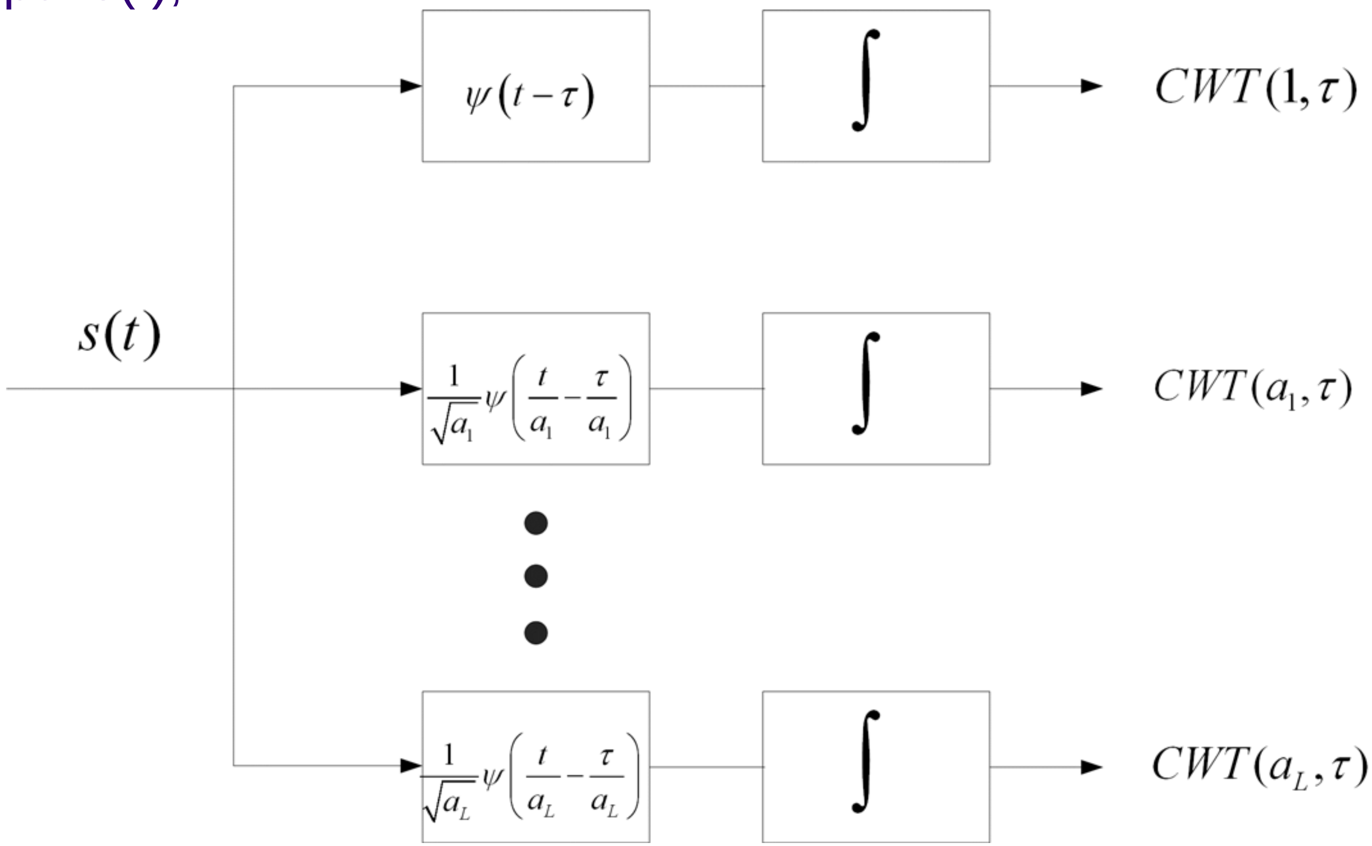
•(iii) it can be also seen to compute the inner product of scaled signal s(at) with $\text{sqrt}(a)\psi(t)$, at shift τ/a.

•(iv) it also follows that CWT is the output of a bandpass filter of impulse response $\text{sqrt}(a)\psi(-t)$ with input *s*(at), at the time instant τ/a.

•These four interpretations give rise to different implementations of the wavelet transform. The choice depends on applications and on the algorithms available. Two cases are important to distinguish:

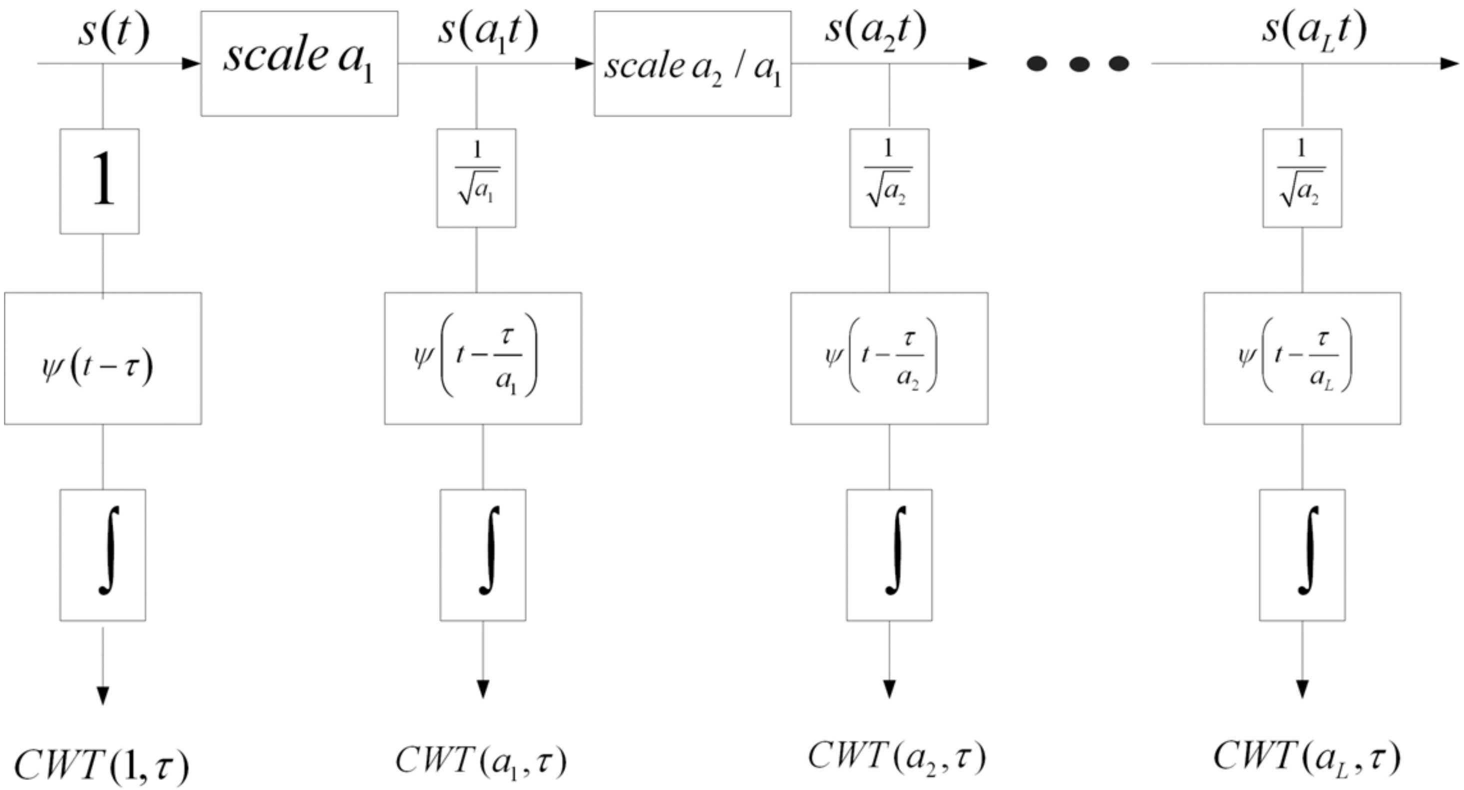
Sub-band decomposition approach[5,6]

1) cross-correlation between *s*(t) and the baby wavelets is equivalent to finding outputs of a bank of bandpass filters of impulse responses ψ(-t/a)/sqrt(a) with input *s*(t),



Sub-band decomposition approach[5,6]

2) successively scaled versions of *s*(t) are passed through identical bandpass filters to give the transform.



By discretizing *s*(t) and scaling it successively by a factor of two implementation of CWT shown above becomes computationally very fast.

Sub-band decomposition approach[5,6]

Four different types of wavelet transform can be distinguished:

(i) The CWT:

$$CWT(a,\tau)=\frac{1}{\sqrt{a}}\int s(t)\psi\left(\frac{t-\tau}{a}\right)dt$$

(ii) The discrete parameter wavelet transform (DPWT):

$$DPWT(m,n)=a_0^{-\frac{m}{2}}\int s(t)\psi\left(a_0^{-m}t-n\tau_0\right)dt$$

where parameters *a* and τ are discretized but the the signal and wavelet function are still continuous. For computational efficiency a_0=2 and τ_0=1 are commonly used.

Sub-band decomposition approach[5,6]

(iii) The discrete time wavelet transform (DTWT):

$$DTWT(m,n) = a_0^{-\frac{m}{2}} \sum_k s(k)\psi\left(a_0^{-m}k - n\tau_0\right)$$

that is time discrete version of DPWT.

(iv) The discrete wavelet transform (DWT):

$$DWT(m,n) = 2^{-\frac{m}{2}} \sum_k s(k)\psi\left(2^{-m}k - n\right)$$

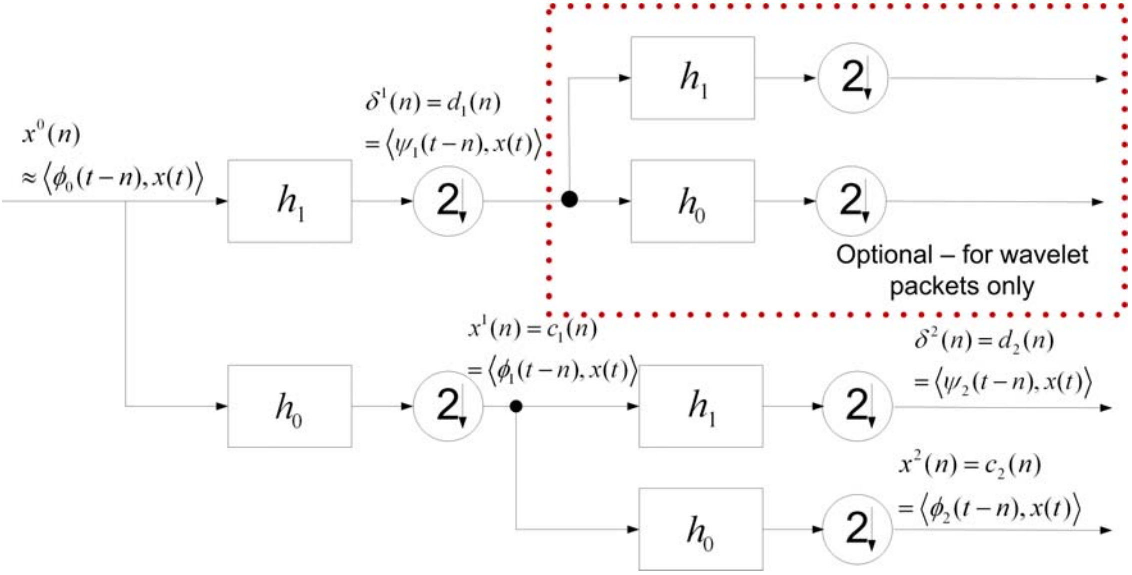
where discrete wavelet $\psi(k)$ can be, but not necessarily, a sampled version of a continuous counterpart.

Sub-band decomposition approach[5,6]

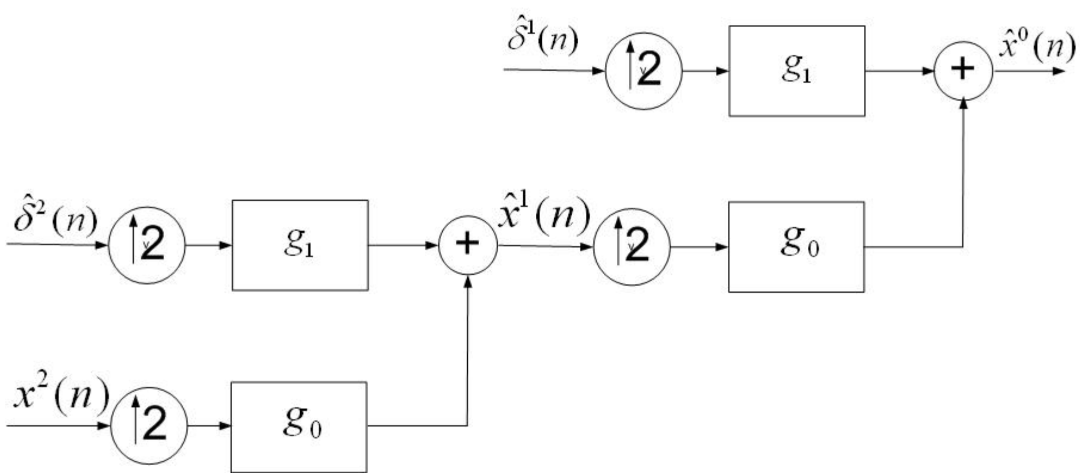
MATLAB implementation of the DWT for *N* discrete resolution levels is obtained by command `swt`. The command `dwt` implements DWT for one fixed resolution level.

DWT (or DPWT) is used to implement the concept of MRA i.e. decomposing signal *s*(k), or *s*(t), successively into coarse and details coefficients, where Impulse responses of the LP and HP filters are wavelet dependent. For the Haar's wavelet LP=([1/2 1/2] and HP=[-1/2 -1/2].

Sub-band decomposition approach[5,6]



Wavelet transform: signal analysis.



Inverse wavelet transform: signal synthesis.

Sub-band decomposition approach[5,6]

Decimation by 2 at each stage is not necessary, in which case we have non-decimated wavelet transform. However, it saves memory and enables faster implementation.

The potential potential problem with decimation is that signals must be bandlimited in order to avoid aliasing (spectra overlapping) problem. In some applications such as speech coding (compression) signal is decomposed on one side and synthesized on another side. In this case it possible to design analysis (decomposition) and synthesis filters that satisfy perfect reconstruction conditions, in which case error caused by aliasing in decomposition stage is canceled out synthesis stage.

In DCA we only decompose signals to locate subband where sources are least dependent. Thus, any error caused by aliasing can not be cancelled out.

Sub-band decomposition approach[5,6]

Two solutions for the aliasing problem are:

i) use nondecimated version of the wavelet transform.

ii) use complex wavelet transform also known dual tree wavelet transform (DTWT)[10,11].

DTWT uses complex signal instead of real one known as analytical continuation. Analytic continuation of real signal *s*(t) is given by

$$\tilde{s}(t) = s(t) + j\hat{s}(t)$$

where

$$\hat{s}(t) = \frac{1}{\pi}\int_{-\infty}^{+\infty}\frac{s(\tau)}{t-\tau}d\tau = H(s(t))$$

[10]N. G. Kingsbury, Complex wavelets for shift invariant analysis and filtering of signals, Appl. Comput. Harmon. Anal. 10 (2001) 234-253.

[11]I. W. Selesnic, R. G. Baraniuk, N.G. Kingsbury, The Dual-Tree Complex Wavelet Transform, IEEE Sig. Proc. Mag. (Nov 2005) 123-

Sub-band decomposition approach[5,6]

i.e. real and imaginary parts of analytic signal form a Hilbert pair. The Hilbert transform can be seen as convolution of *s*(t) with 1/πt. Thus its spectrum is:

$$\hat{S}(\omega) = S(\omega) \times \int_{-\infty}^{+\infty} \frac{1}{\pi t} e^{-j\omega t} dt = S(\omega) \times \left(-jsign(\omega)\right)$$
$$= -jsign(\omega)S(\omega)$$

Spectrum of analytic signal is obtained as:

$$\tilde{S}(\omega) = S(\omega) + j\hat{S}(\omega) = S(\omega) + j\left(-jsign(\omega)S(\omega)\right)$$
$$= S(\omega) + sign(\omega)S(\omega)$$
$$= \begin{cases} 2S(\omega) & \omega > 0 \\ S(0) & \omega = 0 \\ 0 & \omega < 0 \end{cases}$$

Sub-band decomposition approach[5,6]

Thus, for signal $s(t)$ bandlimited in $[-\omega_h, \omega_h]$ the no-overlap condition implies: $\omega_s \geq 2\omega_h$. However, for analytic signal $\tilde{s}(t)$ the no-overlap condition implies: $\omega_s \geq \omega_h$, whereas ω_s is sampling frequency.

Hence, analytic signal can be sampled at the two times smaller frequency for the same amount of error caused by aliasing. That is the motivation for introducing the complex wavelet transform.

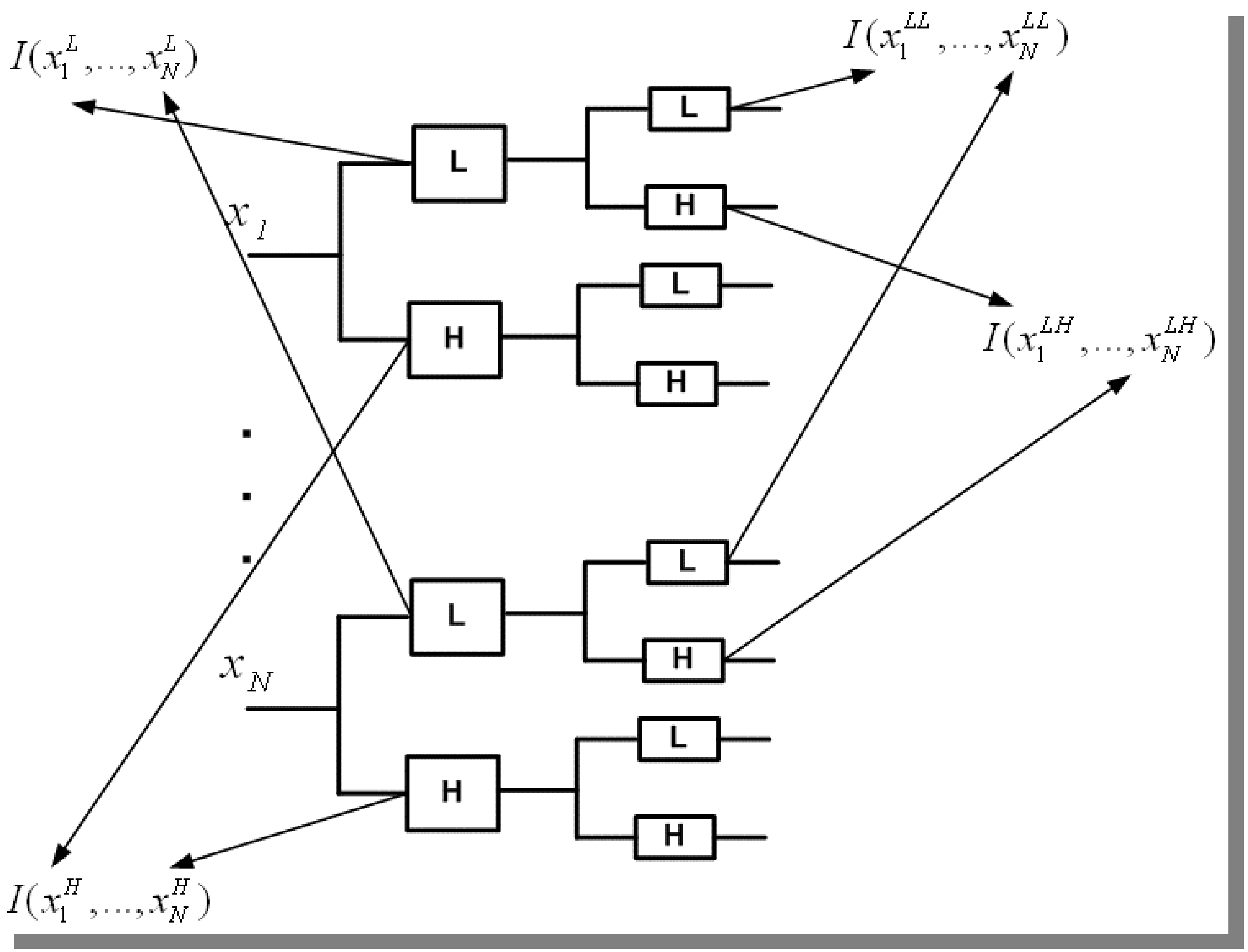
Sub-band decomposition approach[5,6]

Mutiscale analysis SDICA

$$k^* = \arg\min_k I\left(x_1^k, \ldots, x_N^k\right)$$

$$\mathbf{W} \cong ICA\left(\mathbf{x}_k\right)$$

$$\mathbf{s} \cong \mathbf{W}\mathbf{x}$$



Sub-band decomposition approach[5,6]

• In order to locate sub-band with least dependent components we have used *small cumulant approximation* to measure the mutual information[12,5] between the components of the measured signals in the corresponding nodes of the wavelet tree:

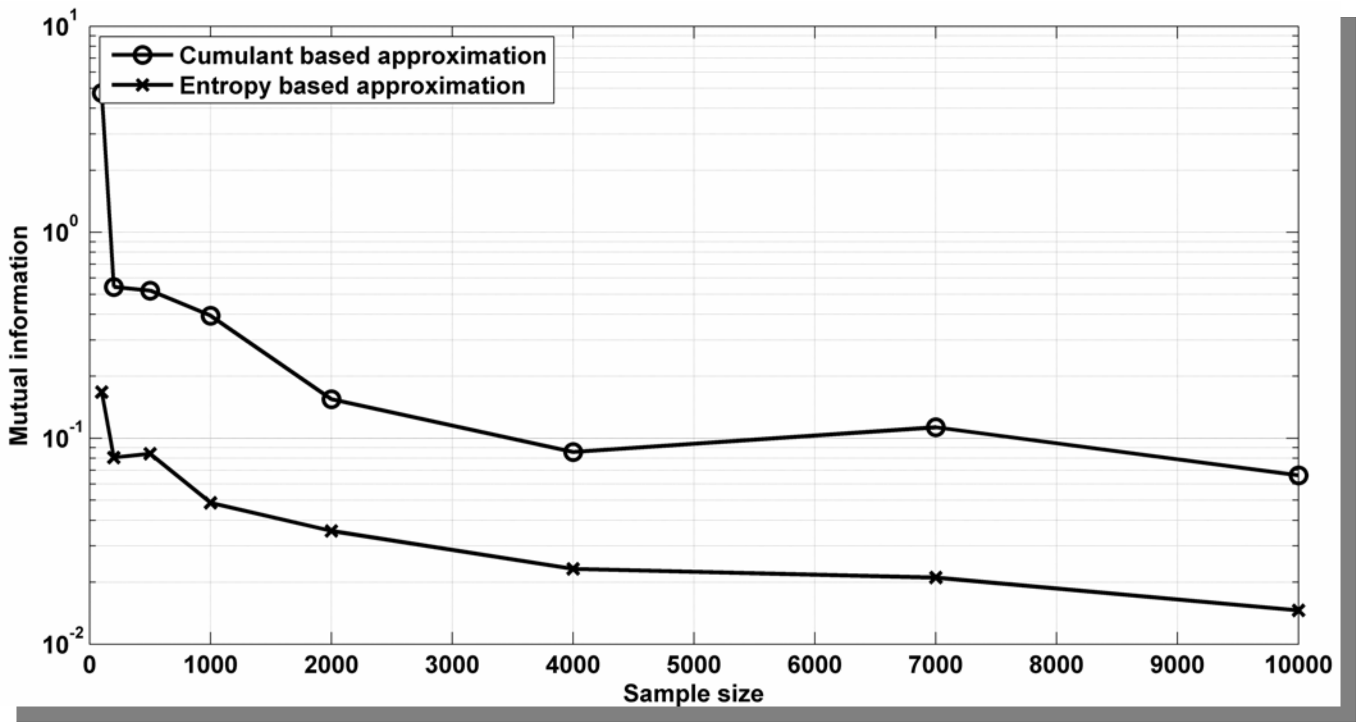
$$\hat{I}_k^j\left(x_{k1}^j, x_{k2}^j, \ldots, x_{kN}^j\right) \approx \frac{1}{4}\sum_{\substack{0\le n<l\le N\\ n\neq l}} cum^2(x_{kn}^j, x_{kl}^j) + \frac{1}{12}\sum_{\substack{0\le n<l\le N\\ n\neq l}}\left(cum^2(x_{kn}^j, x_{kn}^j, x_{kl}^j) + cum^2(x_{kn}^j, x_{kl}^j, x_{kl}^j)\right)$$

$$+\frac{1}{48}\sum_{\substack{0\le n<l\le L\\ n\neq l}}\left(cum^2(x_{kn}^j, x_{kn}^j, x_{kn}^j, x_{kl}^j) + cum^2(x_{kn}^j, x_{kn}^j, x_{kl}^j, x_{kl}^j) + cum^2(x_{kn}^j, x_{kl}^j, x_{kl}^j, x_{kl}^j)\right)$$

where *j* represents scale index and *k* represents sub-band index at the appropriate scale.

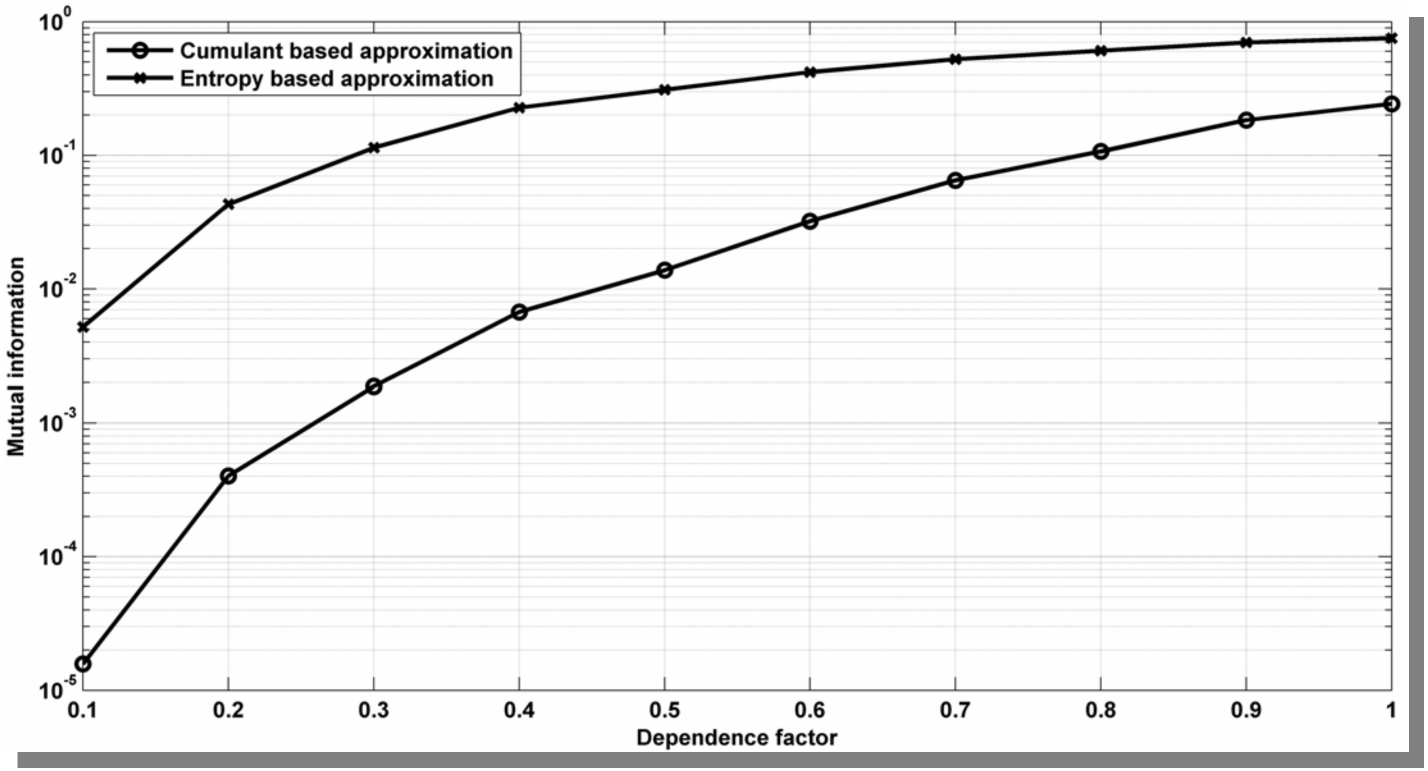
[12] J.F. Cardoso,Dependence, correlation and gaussianity in independent component analysis, J. Mach. Learn. Res. 4 (2003) 1177-1203

Estimation of the Mutual information



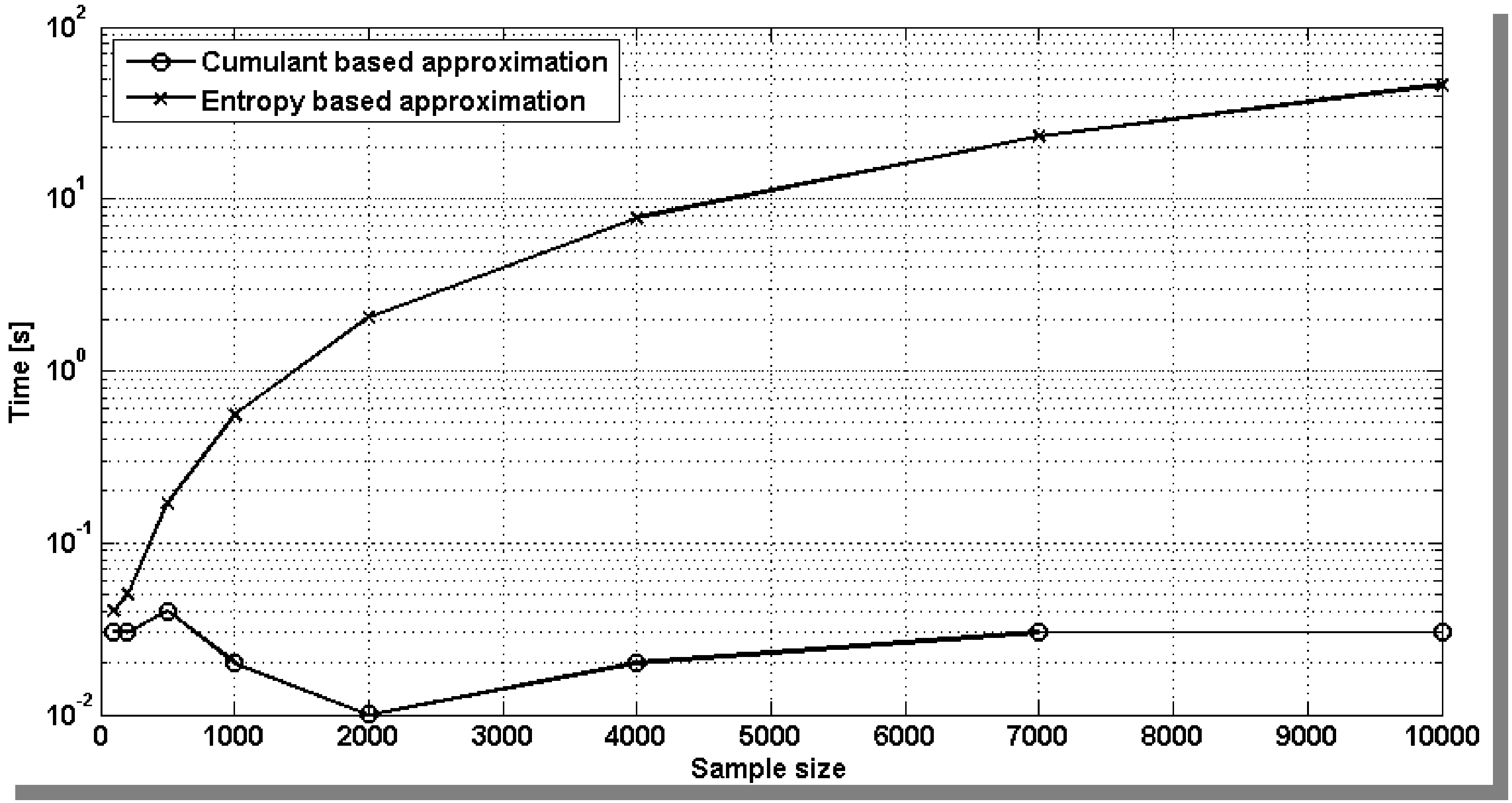
MI approximation as a function of the sample size for two Laplacian distributed processes. ‘x’ denotes entropy-based approximation while ‘o’ denotes cumulant-based approximation.

Estimation of the Mutual information



MI approximation as a function of statistica dependence factor. Two uniformly distributed processes were used as independent processes. Normally distributed process was added as a dependent process with the scale factor that influenced dependence level. ‘x’ denotes entropy-based approximation while ‘o’ denotes cumulant based approximation.

Estimation of the Mutual information



Computation time for two approximations of the MI as a function of the sample size. Two Laplacian distributed processes were used as independent processes. ‘x’ denotes entropy-based approximation. while ‘o’ denotes cumulant-based approximation.

Sub-band decomposition approach[5,6]

•In [5] it has been proven that sub-band index *k* with the least dependent components of the observed signals corresponds with the sub-band index with least dependent components of the source signals.

•The cumulant based approximation is demonstrated in [5] to be consistent estimator of the mutual information an be more computationally efficient than entropy based formulation. Computational complexity of entropy based estimator is caused by necessity to estimate density function, especially multidimensional density function.

$$\hat{I}_H(\mathbf{y}) = \sum_{n=1}^{N} \hat{H}(y_n) - \hat{H}(\mathbf{y})$$

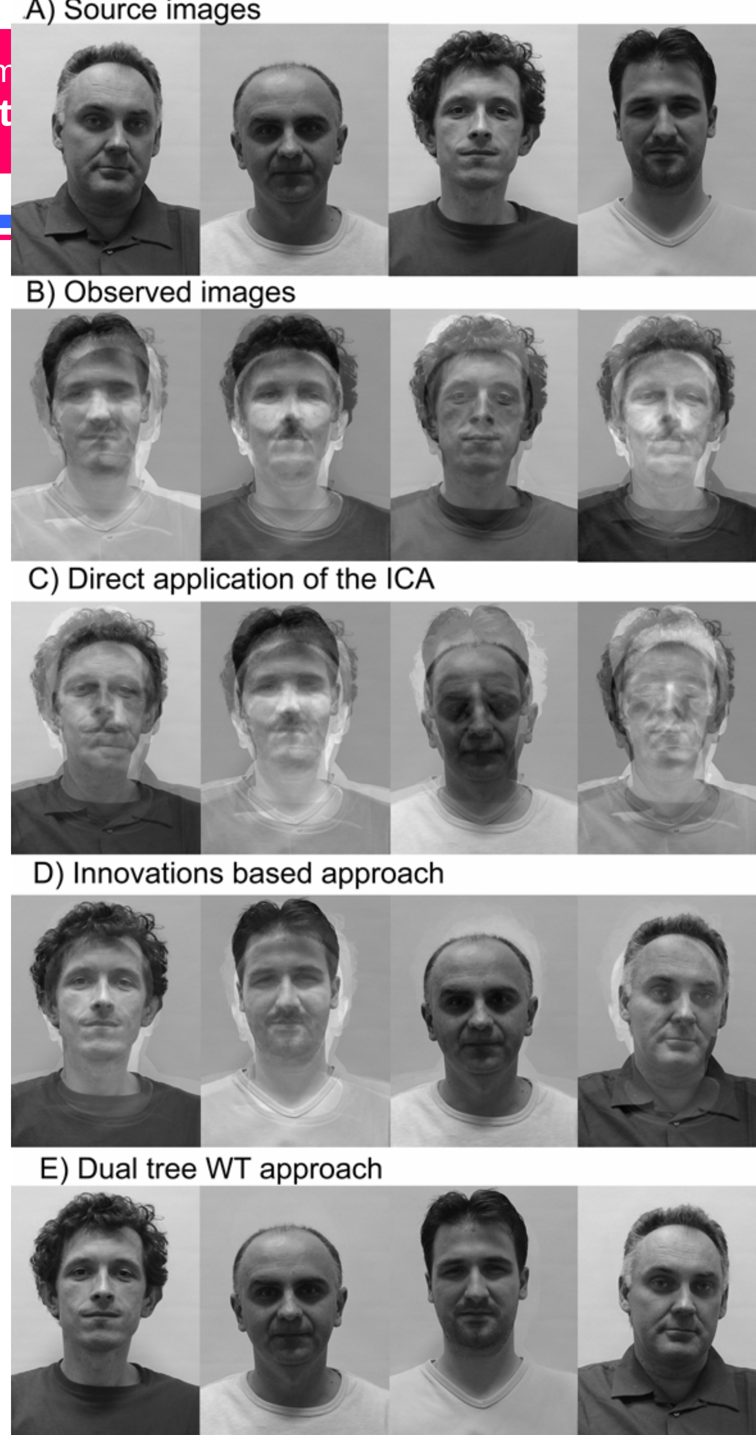
$$\hat{H}(y_n) = -\frac{1}{T}\sum_{t=1}^{T} \log \hat{p}_{y_n}(y_n(t))$$

$$\hat{H}(\mathbf{y}) = -\frac{1}{T}\sum_{t=1}^{T} \log \hat{p}_{\mathbf{y}}\left(\mathbf{y}(t)\right)$$

Separation of images of human faces

• Wavelet packets approach to blind separation of statistically dependent sources has been tested on separation of the images of human faces. They are know to be highly dependent.

• A background Gaussian noise has been added as an wide-band interferer to all source images with an average SNR $\cong$30dB.


A) Source images
B) Observed images
C) Direct application of the ICA
D) Innovations based approach
E) Dual tree WT approach

Single-frame multi-channel blind image deconvolution

• It has been demonstrated in [13] the possibility to formulate "model free" single frame blind image deconvolution as a multichannel BSS problem solved by DCA algorithms.

•Benefit: no *a priori* knowledge about the size and origin of the blurring kernel is required.

$$g(x,y)=\sum_{s=-M}^{M}\sum_{t=-M}^{M}h(s,t)f(x+s,y+t)$$

[13] I. Kopriva, Approach to Blind Image Deconvolution by Multiscale Subband Decomposition and Independent Component Analysis, Journal Optical Society of America A 24 (2007) 973-983.

Single-frame multi-channel blind image deconvolution

The key insight is the Taylor series expansion of *f(x+s,y+t)* around *f(x,y)* first used by Umeyama in [14].

$$f(x+s,y+t)=f(x,y)+sf_x(x,y)+tf_y(x,y)+\dots$$

Then degraded image is obtained as

$$g(x,y)=a_1 f(x,y)+a_2 f_x(x,y)+a_3 f_y(x,y)+\dots$$

where

$$a_1=\sum_{s=-K}^{K}\sum_{t=-K}^{K}h(s,t)\quad a_2=\sum_{s=-K}^{K}\sum_{t=-K}^{K}sh(s,t)\quad a_3=\sum_{s=-K}^{K}\sum_{t=-K}^{K}th(s,t)$$

The PSF coefficients are absorbed into mixing coefficients. No *a priori* knowledge about the nature of the blurring process or size of the blurring kernel is required.

[14] S. Umeyama, Scripta Technica, Electron Comm Jpn, Pt 3, 84(12), 1-9 (2001).

Single-frame multi-channel blind image deconvolution

Implicit Taylor expansion assumes smoothness of the unknown up to certain degree which depends on the strength of blurring process. This implies existence of higher order derivatives which problematic for images with edges. It is, therefore, concluded that proposed concept is efficient for weak blur.

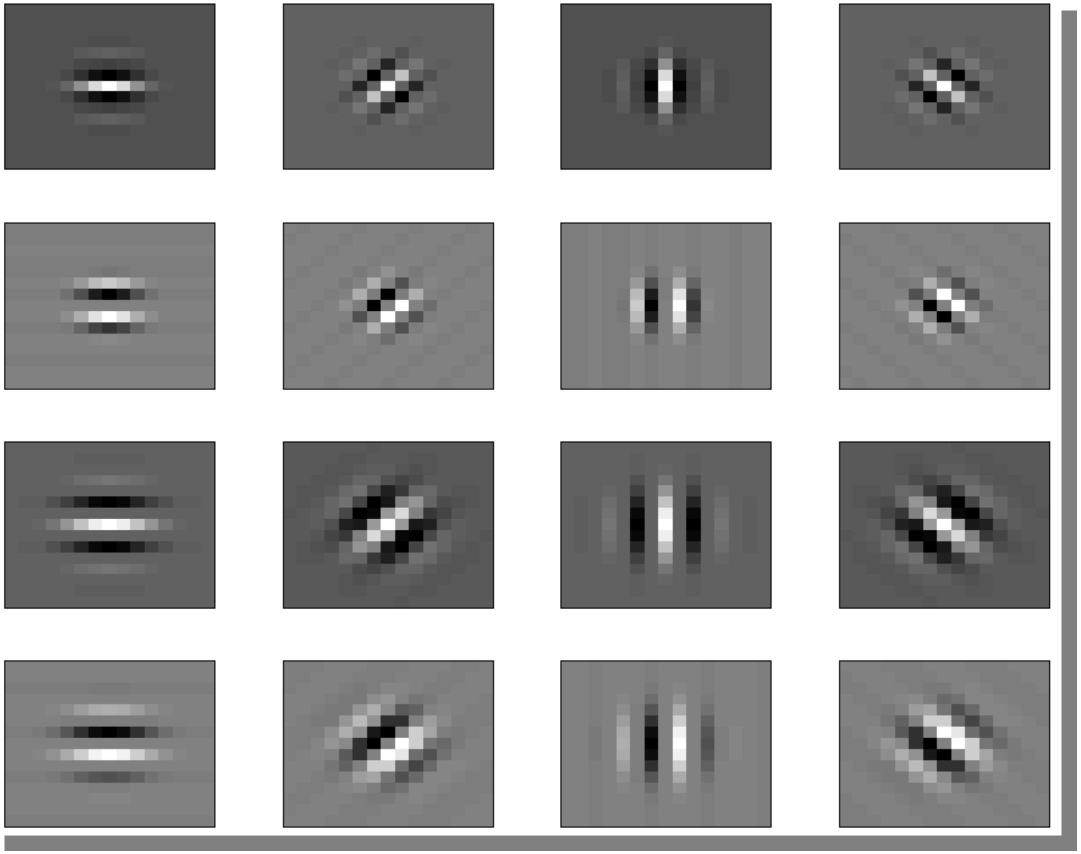
A multi-channel representation is obtained by applying a bank of 2-D Gabor filters to the degraded image *g(x,y)*.

A single frame multi-channel model is obtained as :

$$\mathbf{G} = \begin{bmatrix} \mathbf{g}^T \\ \mathbf{g}_1^T \\ \cdots \\ \mathbf{g}_L^T \end{bmatrix} \cong \begin{bmatrix} a_1 \;\; a_2 \;\; a_3 \ldots \\ a_{11} \; a_{12} \; a_{13} \ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{L1} \; a_{L2} \; a_{L3} \ldots \end{bmatrix} \begin{bmatrix} \mathbf{f}^T \\ \mathbf{f}_\mathbf{x}^T \\ \mathbf{f}_\mathbf{y}^T \\ \ldots. \end{bmatrix} = \mathbf{AF}$$

Single-frame multi-channel blind image deconvolution

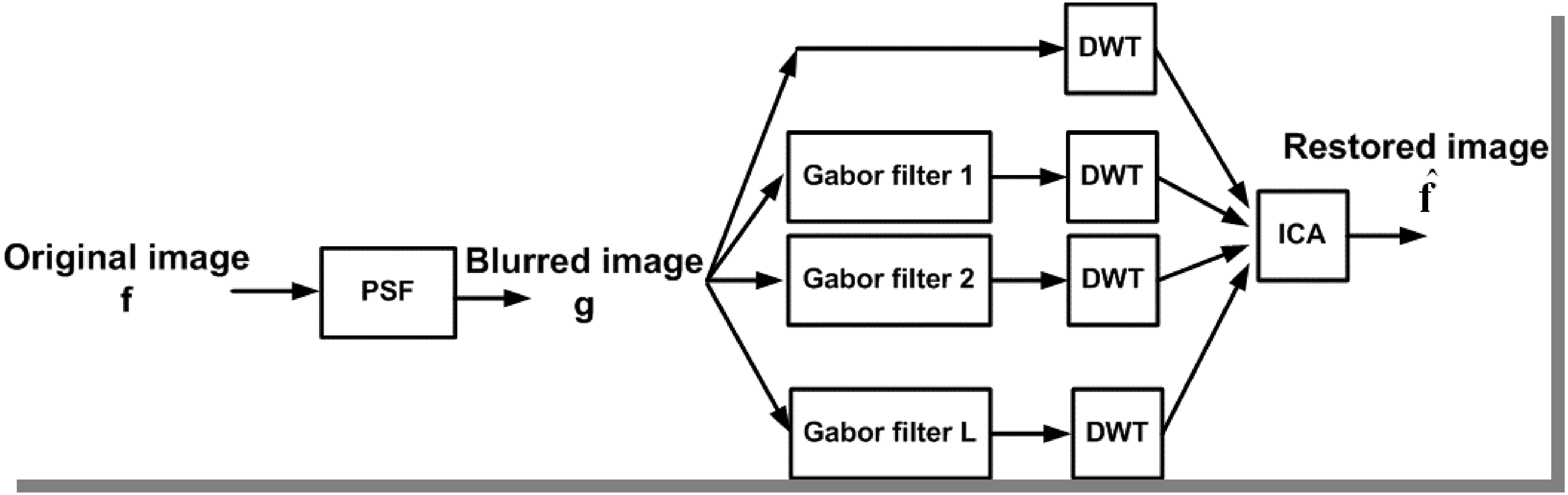
A Gabor filter bank of 7x7 filters with two spatial frequencies and four orientations.



Single-frame multi-channel blind image deconvolution

A problem with the ICA approach arises if the images **f**, $\mathbf{f}_x$, $\mathbf{f}_y$, etc. are not statistically independent.

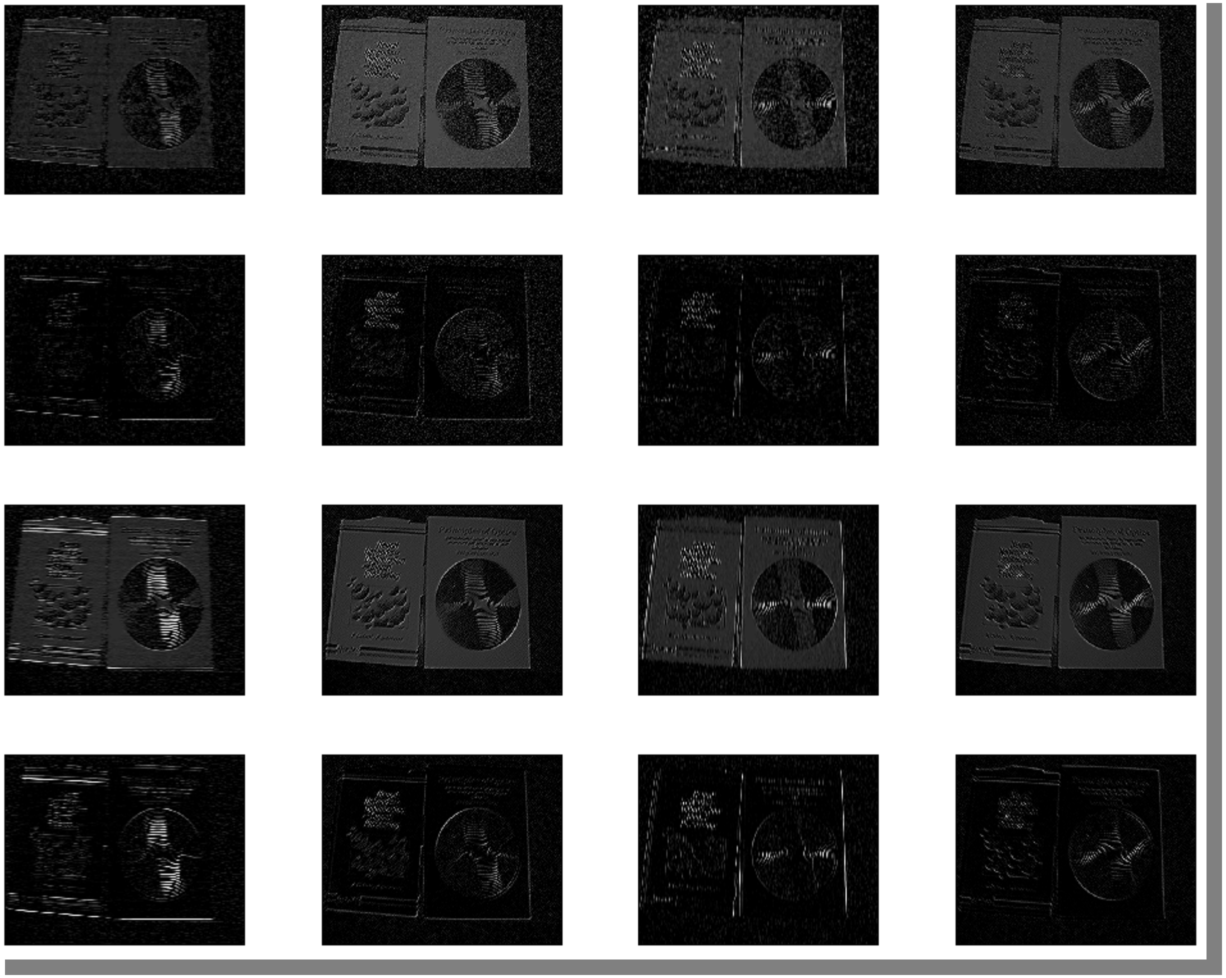
In Ref. [13] DWT based SDICA algorithm has been used to enhance statistical independence among $T(\mathbf{f})$, $T(\mathbf{f}_x)$, $T(\mathbf{f}_y)$, etc.



Single-frame multi-channel blind image deconvolution

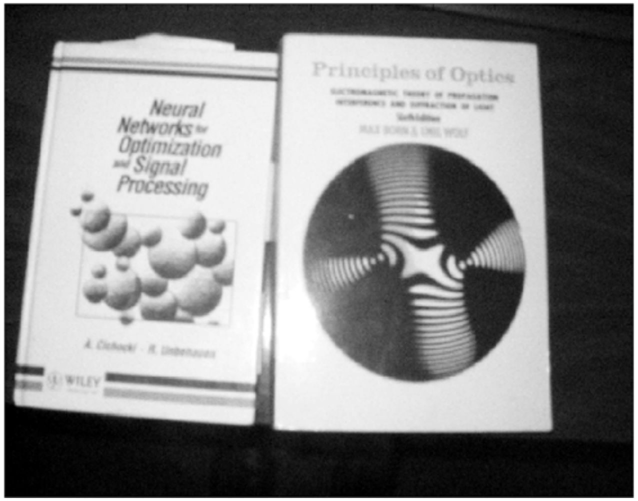
Picture acquired by digital camera in manually de-focused mode.

Output of Gabor filter bank:

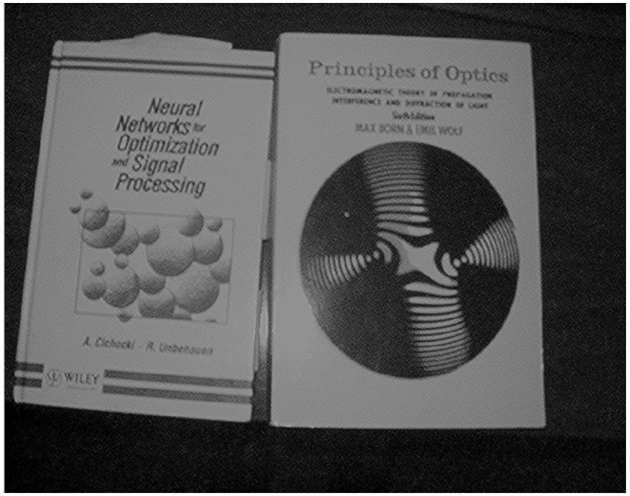


Single-frame multi-channel blind image deconvolution

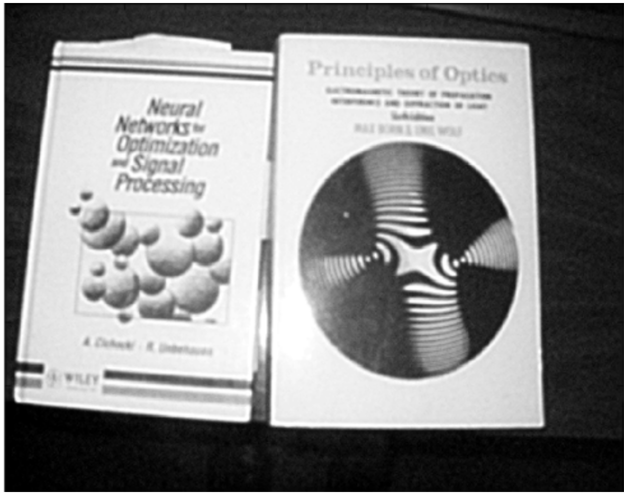
Picture acquired by digital camera in manually de-focused mode.



Blurred



DWT SDICA restored



Blind RL, R=3
(MATLAB command: `deconvblind`)

Blind speech deconvolution[15]

Single channel recording:

$$x(t) = \sum_{\tau=0}^{T} h(\tau)s(t-\tau)$$

The original signal *s*(*t*-τ) can be approximated by Taylor series expansion around *s*(*t*) giving:

$$s(t-\tau) = \sum_{n=0}^{N} \frac{(-\tau)^n}{n!} \frac{d^n s(t)}{dt^n} + \text{H.O.T.}$$

this gives:

$$x(t) \cong \sum_{n=0}^{N} a_{1(n+1)} \frac{d^n s(t)}{dt^n}$$

where $a_{11} = \sum_{\tau=0}^{T} h(\tau)$ $a_{12} = -\sum_{\tau=0}^{T} \tau h(\tau)$ $a_{13} = \sum_{\tau=0}^{T} \left(\tau^2/2\right) h(\tau)$

[15] I. Kopriva, Blind Signal Deconvolution as an Instantaneous Blind Separation of Statistically Dependent Sources, LNCS 4666 (2007) 504-511.

Blind speech deconvolution

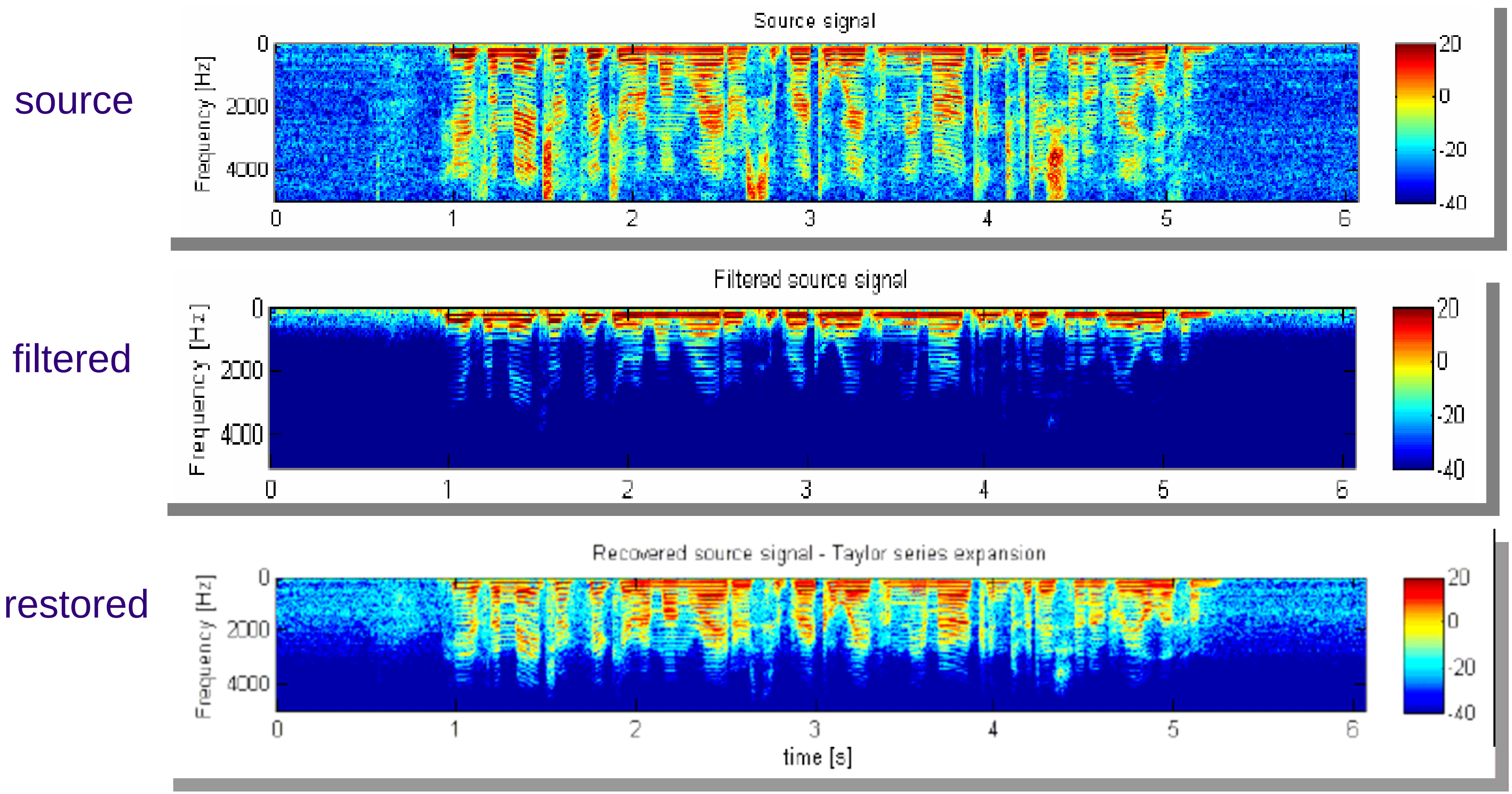
Non-i.i.d. signal (female speech) has been passed through band-limited (low-pass) channel.

2^{nd} order Butterworth filter has been used to model the channel.

1D DWT was used to convert single channel to “multi-channel” representation.

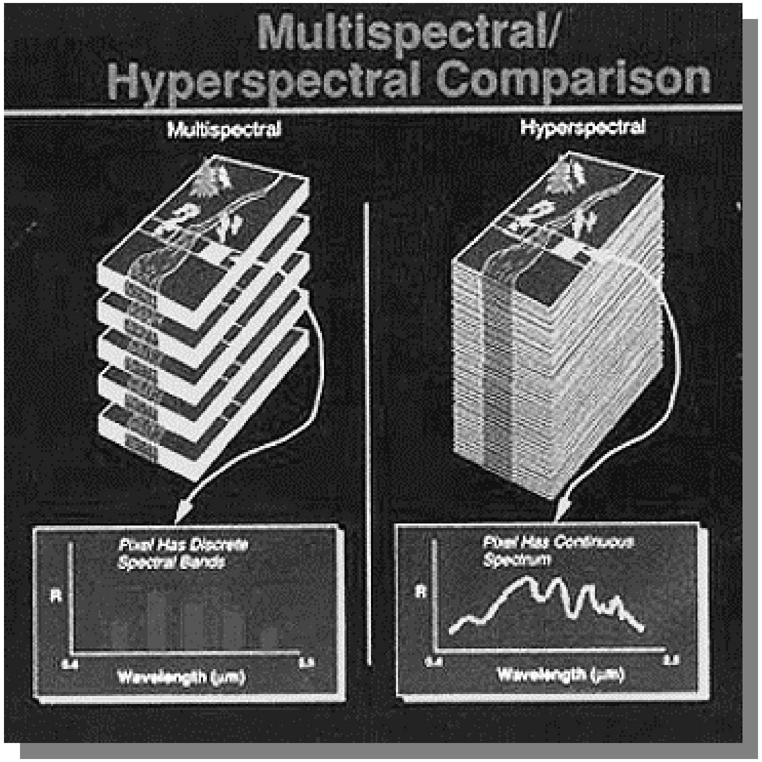
Innovations were used to enhance statistical independence among the time derivatives of the source signal. Prediction-error filter of the 10^{th} order has been used to implement innovations.

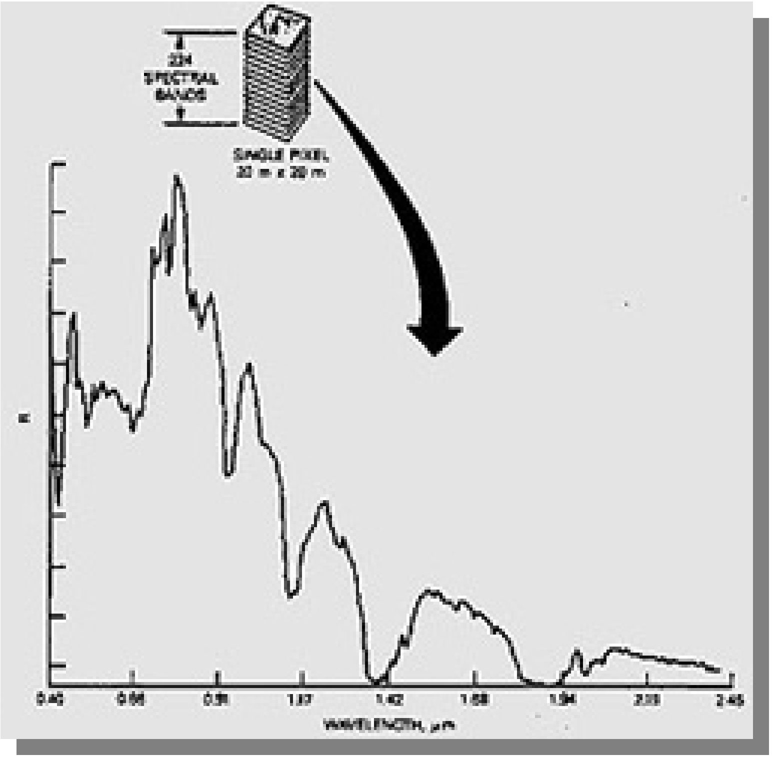
Blind speech deconvolution



Normalized correlation coefficients between source-filtered and restored signals: 0.71774 and 0.88658.

Blind segmentation of multispectral image of the skin tumor[16]





- ❑ Objects with very similar reflectance spectra are *difficult to discriminate*.

[16] I. Kopriva, A. Persin, N. Puizina-Ivic, L. Miric, Robust demarcation of basal cell carcinoma by dependent component analysis based segmentation of multi-spectral fluorescence images, *Journal of Photochemistry and Photobiology B: Biology,* (2007) doi:10.1016/j.jphotobiol.2010.03.013

Hyperspectral/multispectral image and linear mixture model

For an image consisting of N bands and M materials linear mixture model of data is assumed:

$$\mathbf{X} = \mathbf{A}\mathbf{S} = \sum_{m=1}^{M} \mathbf{a}_m \mathbf{s}_m$$

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \ldots & \mathbf{a}_M \end{bmatrix} \equiv \mathbf{A}$$

$$\begin{bmatrix} \mathbf{s}_1 & \mathbf{s}_2 & \ldots & \mathbf{s}_M \end{bmatrix}^T \equiv \mathbf{S}$$

X - measured data intensity matrix, $\mathbf{X} \in \mathbb{R}_{0+}^{N \times T}$

S - unknown class matrix. $\mathbf{S} \in \mathbb{R}_{0+}^{M \times T}$ Each row represents spatial distribution of the class.

A – unknown spectral reflectance matrix. $\mathbf{A} \in \mathbb{R}_{0+}^{N \times M}$ Each column represents spectral profile of the class

Blind segmentation of multispectral image of the skin tumor

Spectral similarity between the sources s_m and s_n implies that corresponding column vectors are close to collinear i.e. $\mathbf{a}_m \cong c\mathbf{a}_n$.

Contribution at certain pixel location t is: $\mathbf{a}_m s_{mt} + \mathbf{a}_n s_{nt} \cong c\mathbf{a}_n s_{mt} + \mathbf{a}_n s_{nt}$. This implies that $\mathbf{s}_n$ and $c\mathbf{s}_m$ are indistinguishable i.e. they are statistically dependent.

Thus, spectral similarity between the sources increases ill-conditioning of the inverse image segmentation problem. The basis (mixing) matrix becomes poorly conditioned and sources become more statistically dependent.

Statistical independence assumed by ICA algorithm on linear mixture model is not satisfied for spectrally similar materials.

Blind segmentation of multispectral image of the skin tumor

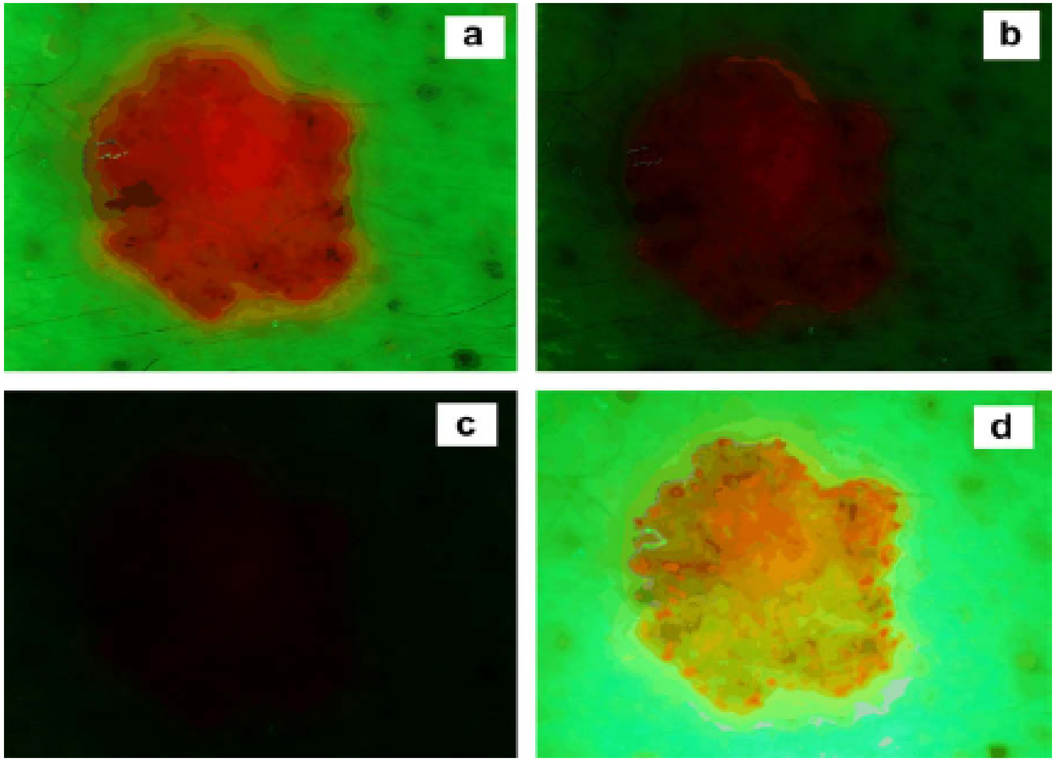


Fig. 1. RGB fluorescent images of the BCC from the first patient acquired under different intensities of illumination: (a) illumination with the maximal intensity I_0; (b) illumination with the intensity $I_0/9.15$; illumination with the intensity $I_0/73.47$; (d) RGB fluorescent image with demarcation line manually marked by the red dots.

Blind segmentation of multispectral image of the skin tumor

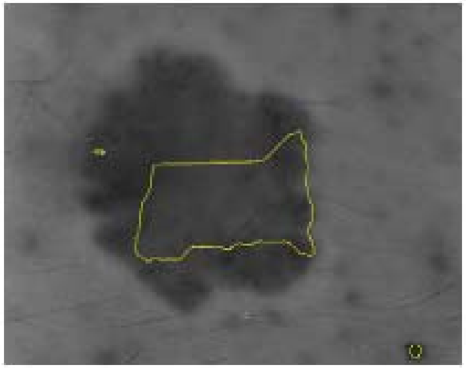


Fig. 3. Evolution curve after 700 iterations calculated by the level set method [46,49] and superimposed on the gray scale version of the fluorescent RGB image shown in Fig. 1a.

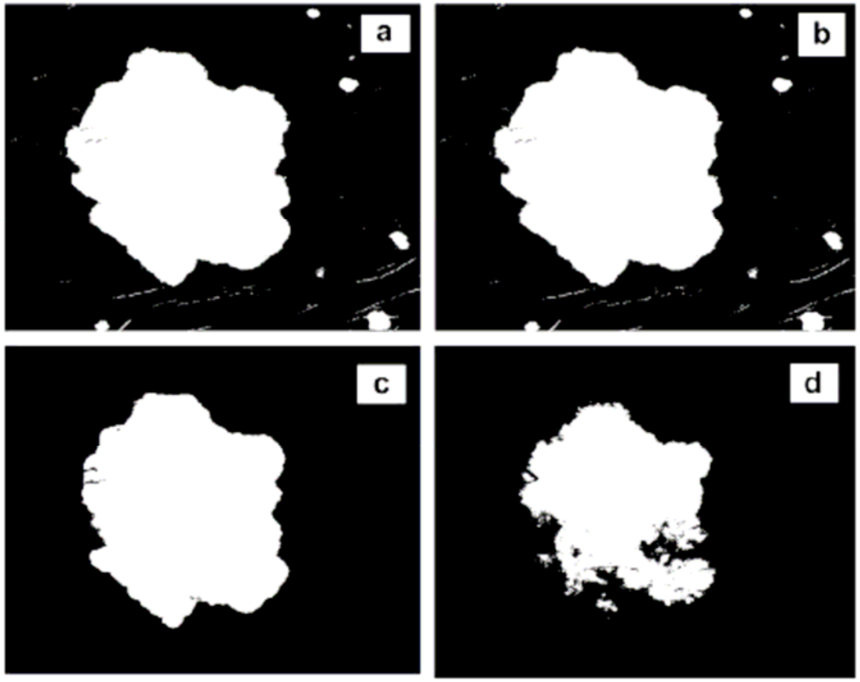


Fig. 4. Segmentation results obtained by: (a) K-means algorithm applied on gray scale version of the Fig. 1a for the interclass distance set to 45; (b) similar as (a) but with interclass distance set to 95; (c) ratio imaging method applied on fluorescent RGB image shown in Fig. 1a with threshold set to 5; (d) similar as (c) but with threshold set to 10.

Blind segmentation of multispectral image of the skin tumor

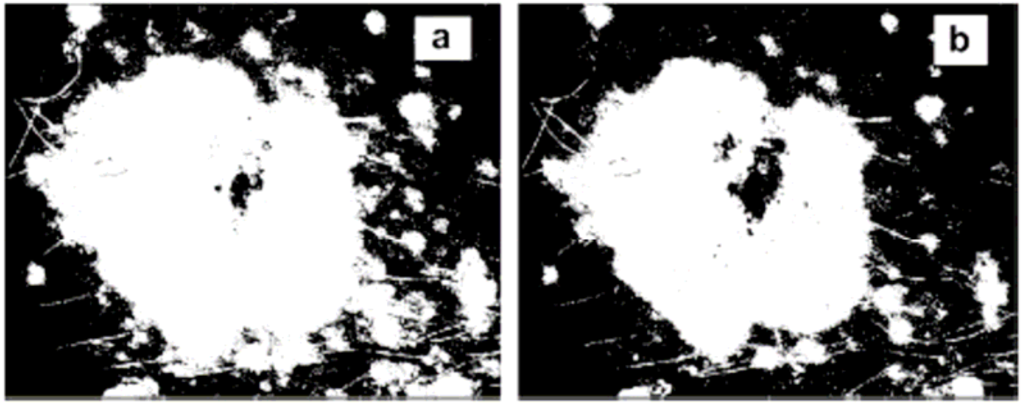


Fig. 5. Segmentation results obtained by K-means algorithm applied on gray scale version of the Fig. 1b for: (a) the interclass distance set to 45 and (b) the interclass distance set to 95.

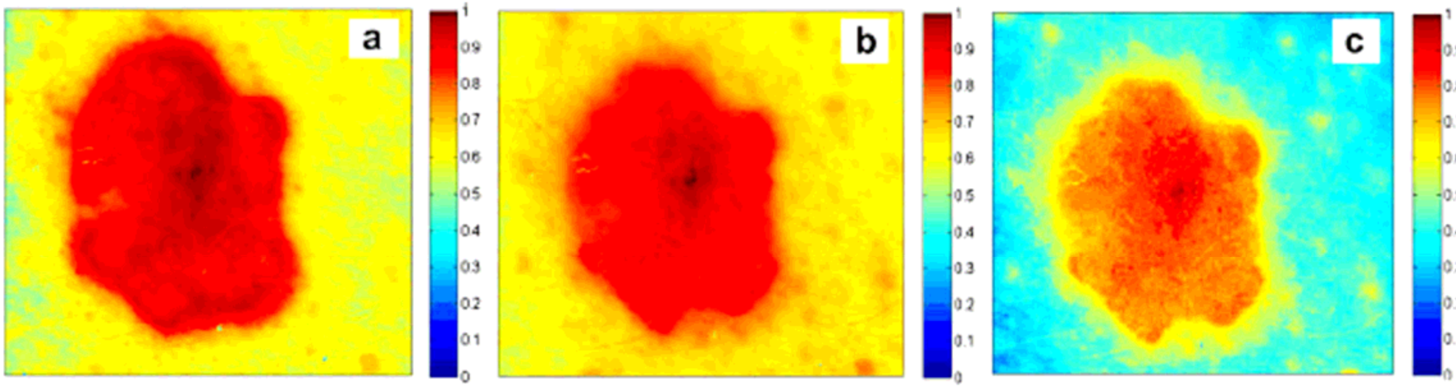


Fig. 6. BCC spatial maps in extracted from fluorescent RGB images shown in Fig. 1a–c by means of EFICA algorithm [36]. Extracted maps are normalized on interval [0, 1] and shown in pseudo-color scale.

Blind segmentation of multispectral image of the skin tumor

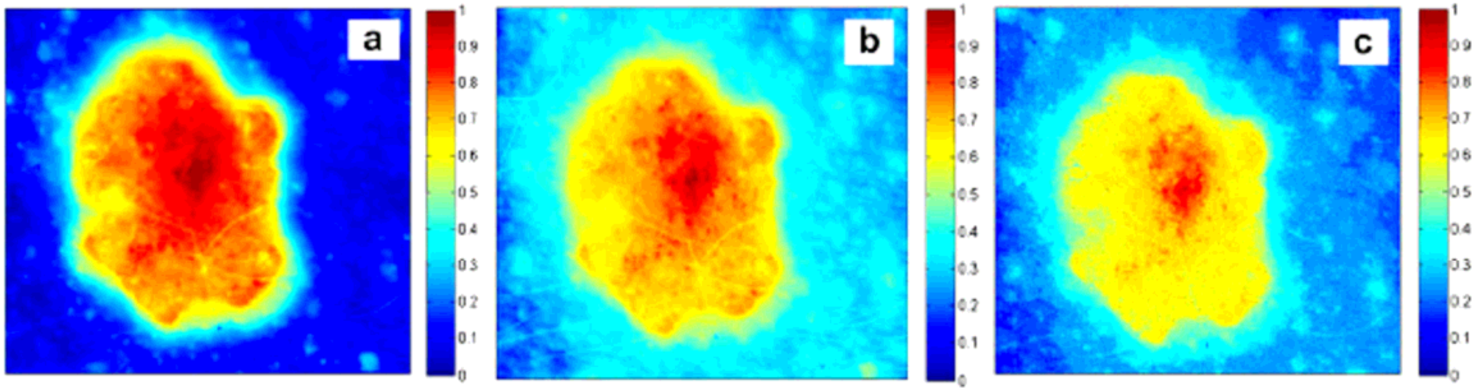


Fig. 7. BCC spatial maps in extracted from fluorescent RGB images shown in Fig. 1a–c by means of DCA-HPF algorithm. Extracted maps are normalized on interval [0, 1] and shown in pseudo-color scale.

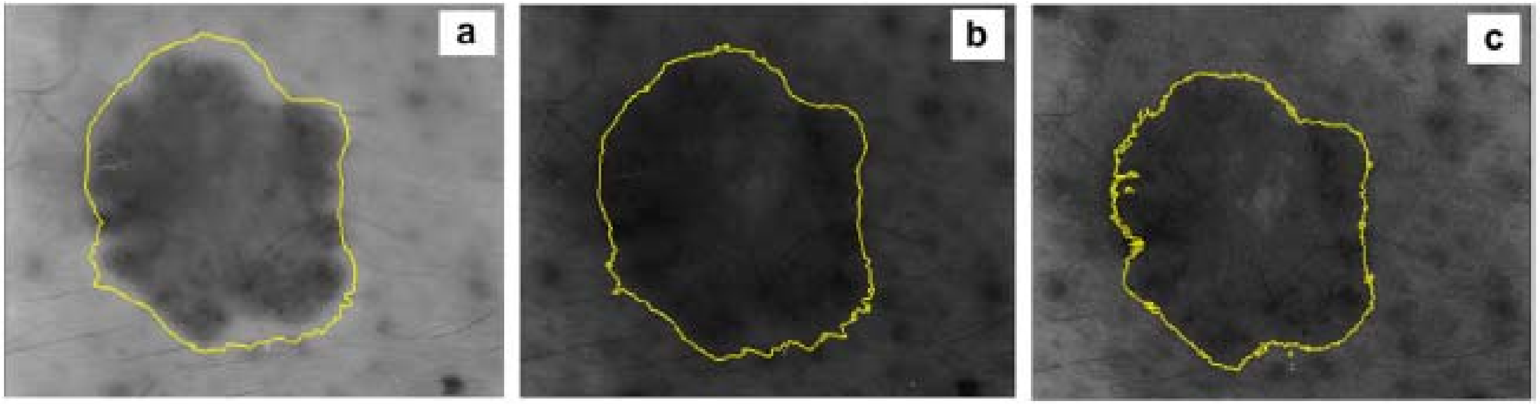


Fig. 8. BCC demarcation lines calculated by means of Canny's edge extraction method from spatial maps shown in Fig. 7a–c, with a fixed threshold set to 0.5. Demarcation lines were superimposed on the gray scale version of the fluorescent RGB images shown in Fig. 1a–c.

Blind segmentation of multispectral image of the skin tumor

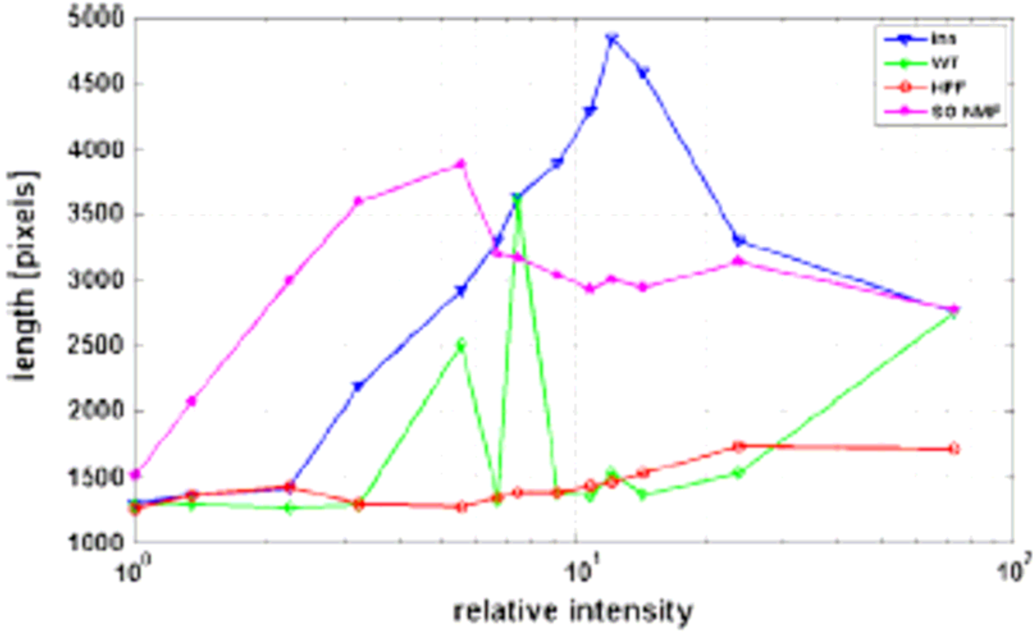


Fig. 9. Estimated lengths of the demarcation lines in pixels. Demarcation lines were calculated by means of Canny's edge extraction method with a fixed threshold set to 0.5, and from the BCC spatial maps extracted from the fluorescent RGB images by means of NMF and DCA algorithms. Legend: red circles – DCA-HPF algorithm [30,31]; green diamonds – DCA WT algorithm [30]; blue triangles – DCA Inn algorithm [28]; magenta stars – SO NMF algorithm [42].

ICA/DCA and NIR Spectroscopy[17]

- NIR is used widely in agriculture, food and beverage industries. Because it is noninvasive it is used increasingly in the medical field.

- NIR spectra of a mixture are often the linear combination of the spectra of its constituent species. It would be useful to extract constituent species from the mixture.

- Classical multivariate linear regression, principal component regression, partial least square regression, as well as Fourier and wavelet transform fail to do that.

- It was shown that ICA (DCA) was able to separate spectra of the constituents from the spectra of their mixtures.

[17]J.Chen and X.Z.Wang, J.Chem. Inf. Comput. Sci. vol 41 2, 992-1001, 2001.

ICA/DCA and NIR Spectroscopy

Component spectra and their mixture:

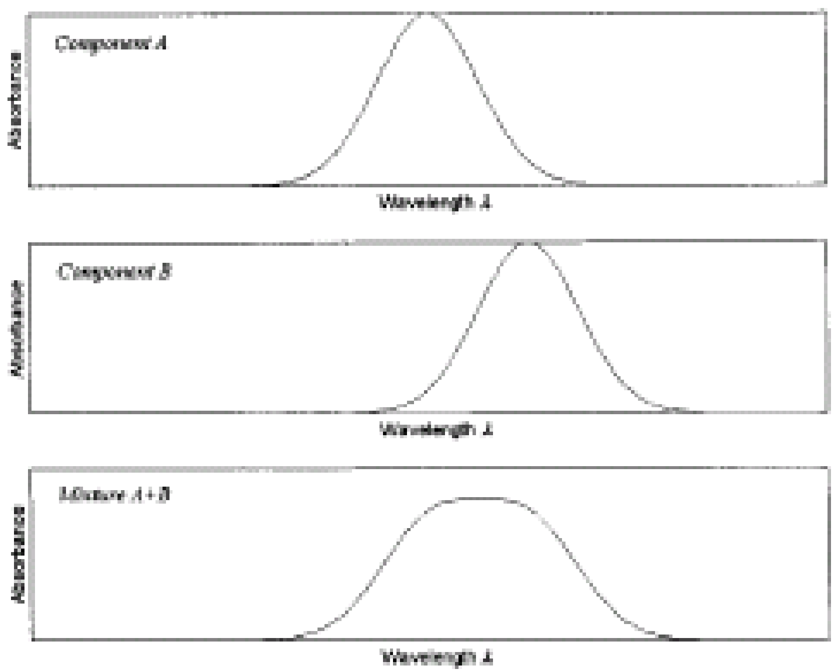


Figure 1. The spectra of *components A* and *B* and their mixture.

$$x_i(\lambda) = c_{i1}s_A(\lambda) + c_{i2}s_B(\lambda)$$

Since component spectra are nonnegative they are **partially correlated** i.e. **statistically dependent**.

ICA/DCA and NIR Spectroscopy

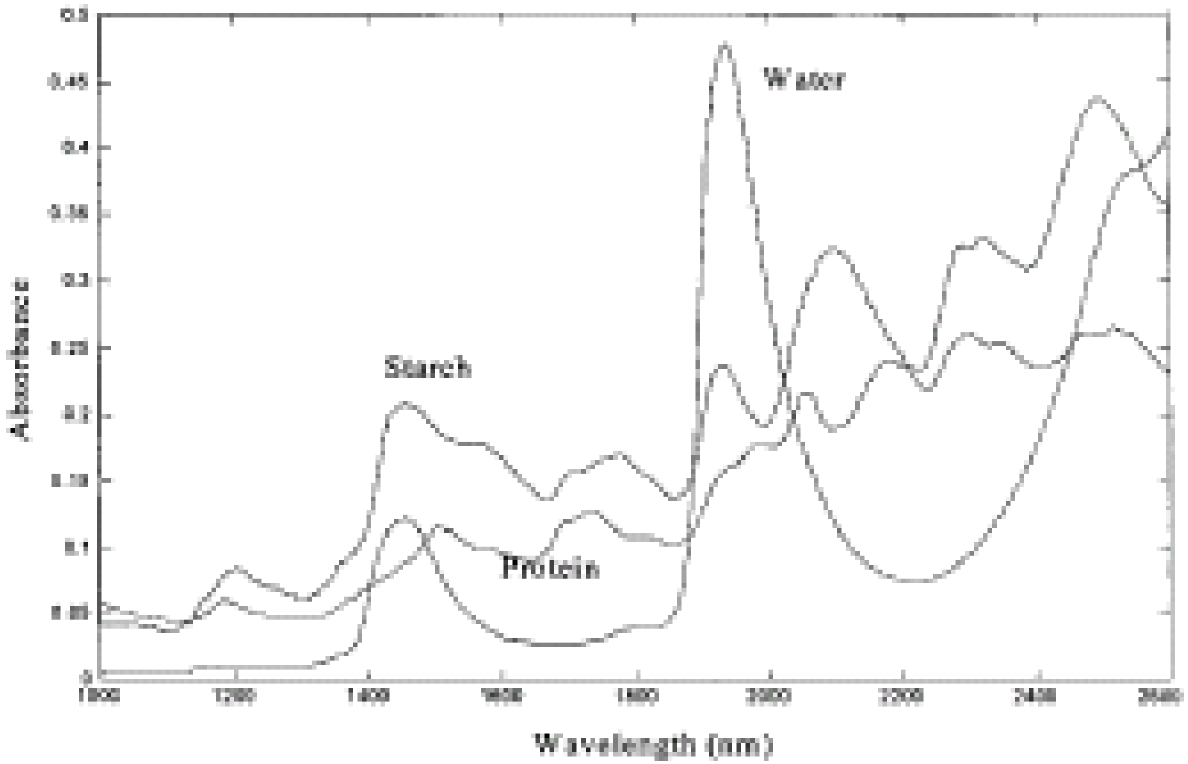


Figure 2. The spectra of water, starch, and protein.

The spectra of water, starch and protein

ICA/DCA and NIR Spectroscopy

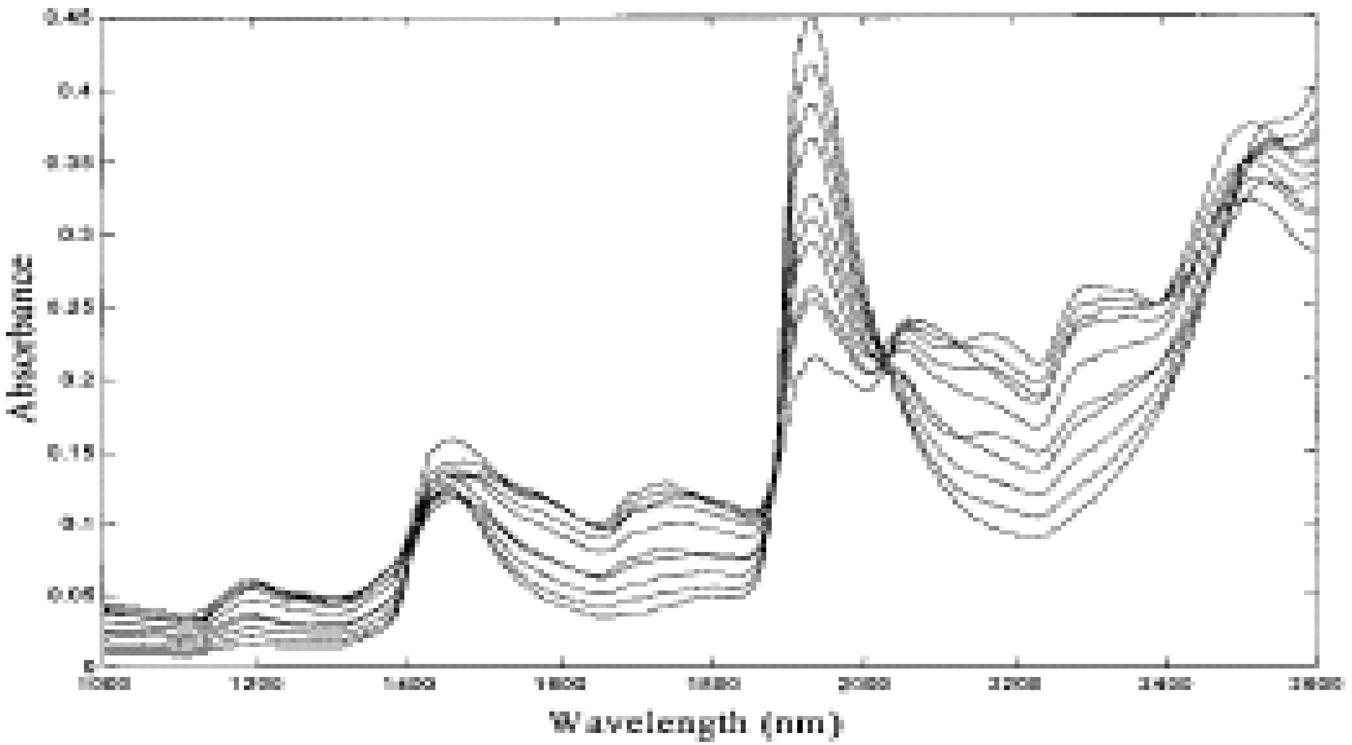
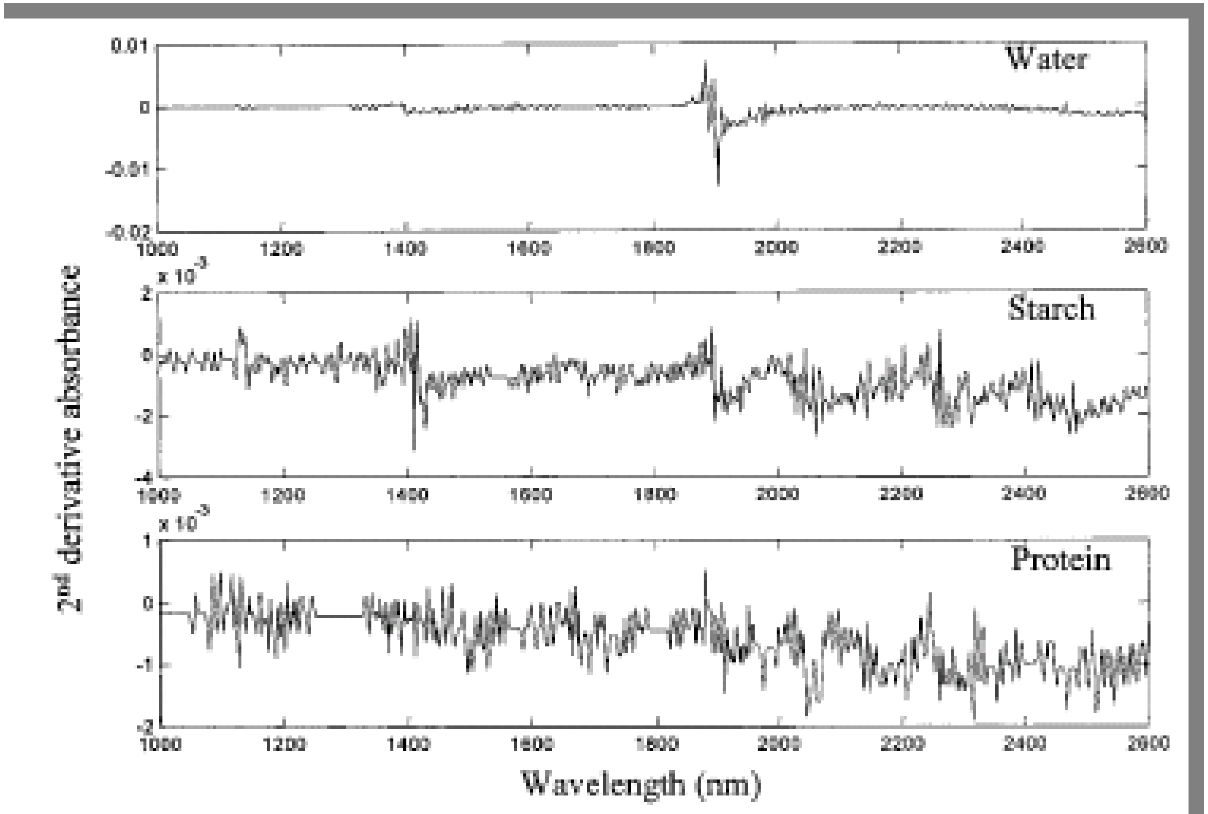


Figure 3. The spectra of 10 samples consisting of water, starch, and protein.

The spectra of mixtures containing different concentrations of water, starch and protein.

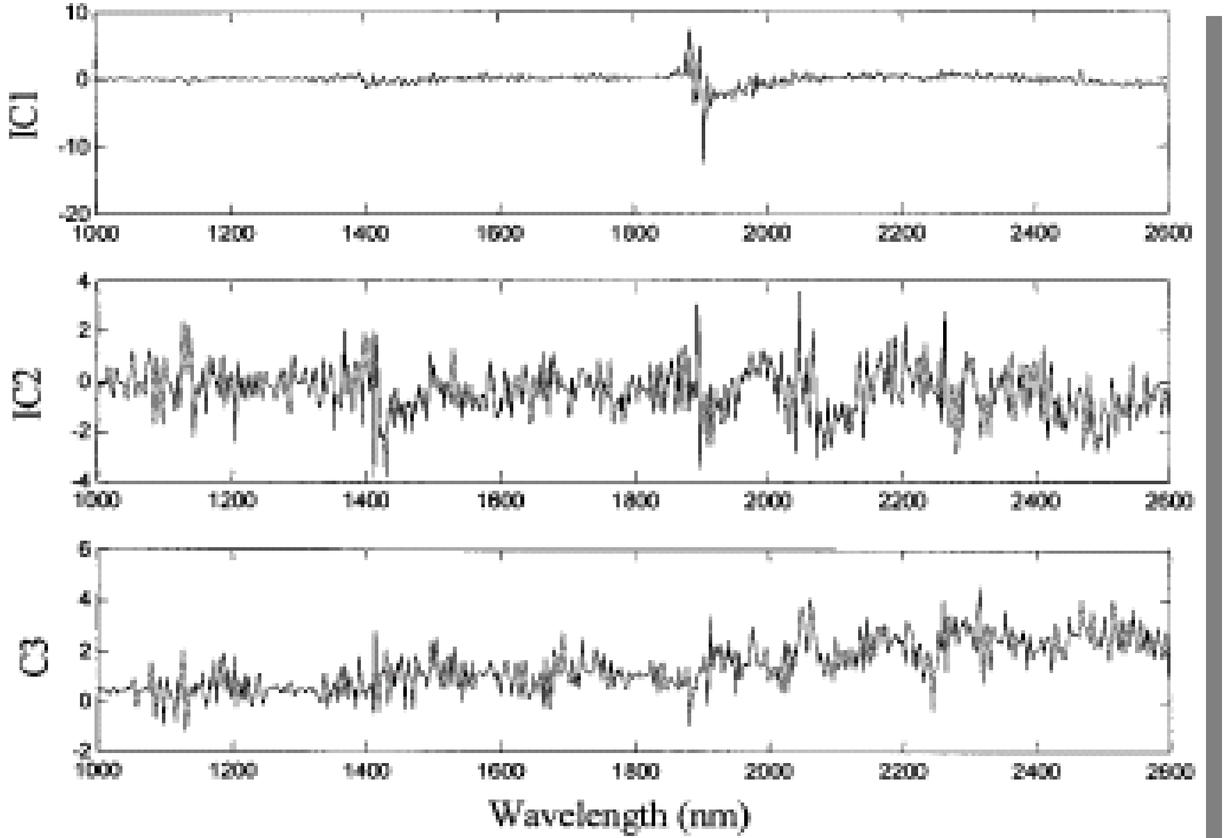
ICA/DCA and NIR Spectroscopy

The NIR spectra need to be the baseline corrected that is done by the 2nd order derivative of the original spectra.



The 2nd order derivative of the original spectra.

ICA/DCA and NIR Spectroscopy



d ICs from Figure 6.

The 2nd order derivative of the ICA/DCA recovered spectra. Correspondence to the original spectra of water, starch and protein is clear.