SPANDIDOS PUBLICATIONS

Blind source separation for the computational analysis of dynamic oncological PET studies

TRIAS THIREOU[1], SOTIRIS PAVLOPOULOS[1], GEORGE KONTAXAKIS[2] and ANDRES SANTOS[2]

[1]Biomedical Engineering Laboratory, National Technical University of Athens, Athens, Greece;
[2]E.T.S.I. de Telecomunicación, Universidad Politécnica de Madrid, Madrid, Spain

Received September 6, 2005; Accepted October 10, 2005

Abstract. The analysis of dynamic positron emission tomography (PET) studies provides clinically useful parametric information, but often requires complex and time-consuming compartmental or non-compartmental techniques. Independent component analysis (ICA), a statistical method used for feature extraction and signal separation, is applied to dynamic PET studies to facilitate the initial interpretation and visual analysis of these large image sequences. ICA produces parametric images, where structures with different kinetic characteristics are assigned opposite values and readily discriminated, improving the identification of lesions and facilitating the posterior detailed kinetic analysis.

Introduction

Although in oncology, the visual inspection of the positron emission tomography (PET) images is practiced routinely for tumor diagnosis, detection of metastases and evaluation of treatment, quantitative measures based on the normalization of tracer concentrations for the injected activity and body weight (standardized uptake values; SUV) are becoming common in the clinical praxis in oncological PET studies (1). However, the use of SUV as a method of classifying tissue areas as benign or malignant is still under discussion among nuclear medicine physicians and oncologists (2,3).

Dynamic 18F-FDG PET studies offer differential diagnostic information and therefore represent an accurate approach to quantify 18F-FDG kinetics. Such studies are increasingly used in oncology for diagnosis, therapy management and evaluation (4). The analysis of dynamic PET sequences, however, often requires complex analysis using compartmental or non-compartmental models, where several difficulties must be overcome, such as determination of the input function of the concentration of the radioactive tracer in plasma, intrinsic inaccuracies at the time of selecting the appropriate compartmental model, or time-consuming computations involving a large amount of image data to be processed.

We investigated the use of independent component analysis (ICA) to reduce the initial amount of image data to a smaller, comprehensive and easily managed set of parametric images. ICA has been shown to produce promising results in the analysis of task-related functional magnetic resonance imaging (fMRI) techniques (5), and extraction of the input function in dynamic myocardial and brain PET studies (6,7).

Materials and methods

Independent component analysis. Independent component analysis (ICA) is a statistical and computational technique for revealing hidden factors that underlie sets of random variables, measurements, or signals.

ICA defines a generative model for the observed multivariate data, which is typically given as a large database of samples. In the model, the data variables are assumed to be linear mixtures of some unknown latent variables, and the mixing system is also unknown. The latent variables are assumed to be non-Gaussian and mutually independent and are called the independent components of the observed data. These independent components, also called sources or factors, can be found by ICA, which is used as a method for blind source separation (8). The intuitive notion of maximum non-Gaussianity or more classical notions like maximum likelihood estimation and minimization of mutual information can be used to derive different objective functions, whose optimization enables the estimation of the ICA model.

Due to its computational and theoretical simplicity, kurtosis, or rather its absolute value, has been widely used as a measure of non-Gaussianity in ICA. However, it can be sensitive to outliers (9) and is not a robust measure of non-Gaussianity. Therefore, other objective functions for ICA estimation have been proposed.

The entropy of a random variable, a basic concept of information theory, can be interpreted as the amount of information given by the observation of a variable. Negentropy is a slightly modified version of differential entropy that is always non-negative, equal to 0 only if a variable has a Gaussian distribution and is invariant for invertible linear

Correspondence to: Dr Sotiris Pavlopoulos, Biomedical Engineering Laboratory, National Technical University of Athens, Athens, Greece
E-mail: spav@biomed.ntua.gr

Key words: independent component analysis, positron emission tomography, standard uptake values

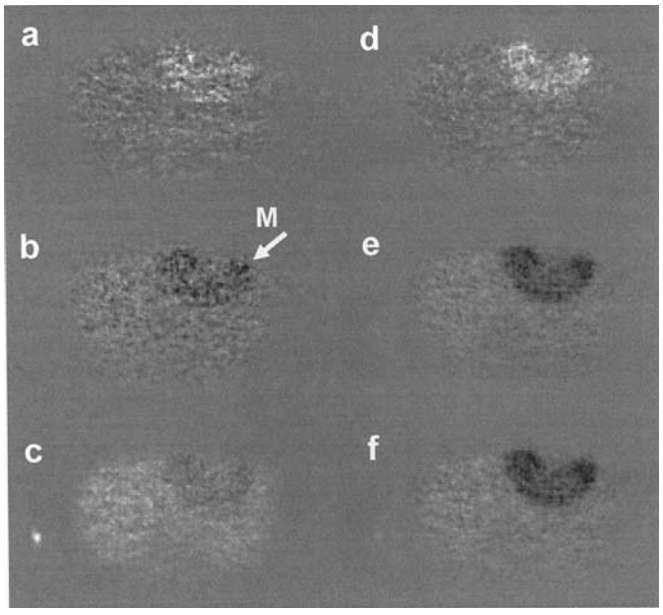

Figure 1. Application of FastICA to a liver metastasis study. The utilized de-correlation approaches are symmetric (a-c) and deflation (d-f). Non-linearity functions are 'pow3' (a and d), 'tanh' (b and e) and 'gauss' (c and f).

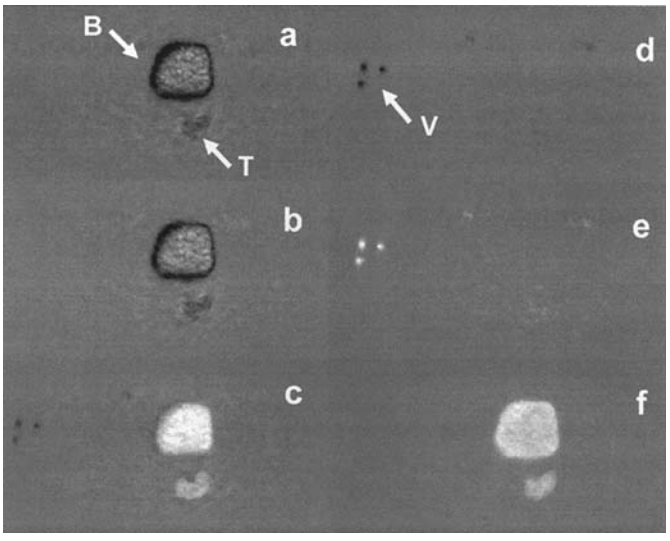

Figure 2. Analysis of a dynamic PET study of a tumor in the vicinity of bladder. Tumor and vessel images generated using FastICA and 'tanh' (a and d), and 'gauss' (b and e) non-linearities with skew-sICA (c) and skew-tICA (f) images of the same study.

transformations. The advantage of using negentropy as a measure of non-Gaussianity is that it is well justified by statistical theory. The problem with using negentropy, however, is that it is computationally difficult. Estimating negentropy using the definition would require an estimate (possibly non-parametric) of the probability density function. Therefore, simpler approximations of negentropy based on the maximum-entropy principle were developed (10) and are used together with a fixed-point (11,12) iteration scheme by the FastICA algorithm (8).

The fixed-point algorithm directly finds independent components of (practically) any non-Gaussian distribution using any non-linear function g(u). Examples include: i) 'pow3,' $g(u) = u^3$; ii) 'tanh,' $g(u) = tanh(a_1 \cdot u)$; and iii) 'gauss,' $g(u) = u \cdot exp(-a_2 \cdot u^2/2)$ where a_1 and a_1 are some suitable constants, often taken as a_1=1 and a_1=1.

To estimate several independent components, the one-unit FastICA algorithm is run using several units ('unit' refers to computational unit, which eventually becomes an artificial neuron). To prevent individual weight vectors of the neurons from converging to the same maxima, the outputs should be de-correlated after every iteration step. The utilized de-correlation approaches are symmetric ('symm'), i.e. estimation of all independent components in parallel, and deflation ('defl'), i.e. estimation of independent components one-by-one.

FastICA has a number of desirable properties including fast convergence confirmed by simulations and experiments on real data (13), selection of a suitable nonlinear function g to obtain robust and/or minimum variance algorithms and computational simplicity. It has been used for separating artifacts in magneto-encephalography data (14), finding hidden factors in financial data (15), reducing noise in natural images (16) and separating signals from interfering ones in telecommunications (17).

Another approach for ICA estimation was derived from a neural network viewpoint (18,19), and is based on maximizing the output entropy (or information flow) of a neural network with non-linear outputs. The principle of network entropy maximization, or infomax, was proven to be equivalent to maximum likelihood estimation. This equivalence requires that the non-linear functions used in the neural network are chosen as the cumulative distribution functions corresponding to the densities of source signals.

Using the infomax approach (18) and a conjugate gradient method (20), the previously developed code is able to execute spatial (sICA) or temporal (tICA) ICA of 2D images (5). Spatial ICA (sICA) seeks a set of mutually independent component (IC) source images and corresponding set of unconstrained time courses. Temporal ICA (tICA) seeks a set of IC source time courses and corresponding set of unconstrained images.

Both sICA and tICA are based on the assumption that the probability density function (pdf) of the independent sources is highly kyrtotic and symmetric. Since this assumption is not warranted for dynamic PET datasets, we proposed to apply skew-ICA to the dynamic PET datasets. Skew-ICA embodies the assumption that images are characterized by the skewness (rather than kurtosis) of their pdfs; an assumption consistent with spatially localized regions of activity (21).

Before applying an ICA algorithm to the data, some preprocessing techniques are commonly used to make the problem of ICA estimation simpler and better conditioned (8). The most basic and necessary preprocessing is to center the data, i.e. make them a zero-mean variable, which implies that source signals are zero-mean as well. This preprocessing is performed solely to simplify the ICA algorithms.

Another useful preprocessing strategy in ICA is to whiten the observed variables, which linearly transforms the observed vector x so its components are uncorrelated and their variances are unified. One popular method of whitening is using the eigenvalue decomposition (EVD) of the covariance matrix. The utility of whitening relies on the number of estimated parameters being reduced. Since whitening is a simpler and more standard procedure than any of the ICA algorithms, it

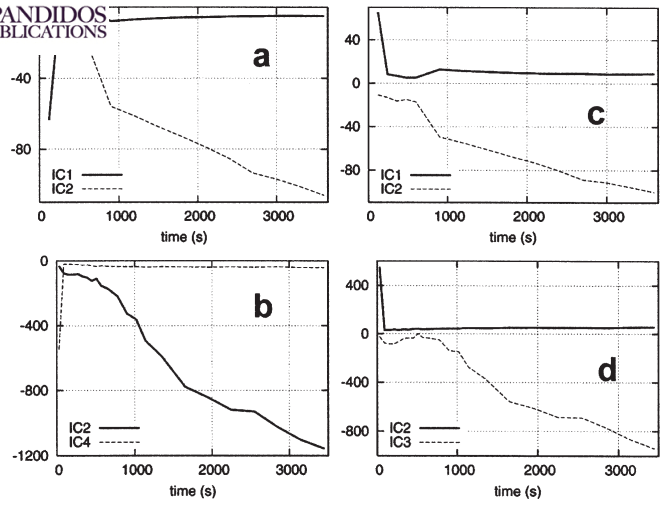

Figure 3. Independent components (ICs) of the studies presented in Figs. 1 (a and c) and 2 (b and d), calculated using the 'tanh'(a and b) and 'gauss' (c and d) functions.

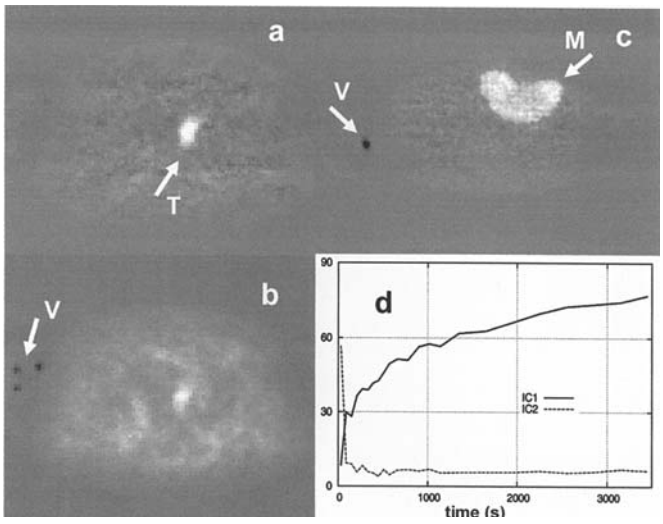

Figure 4. First skew-sICA and skew-tICA image (a and b) of a colorectal tumor recurrence study. First Skew-sICA image of the liver metastasis study presented is Fig. 1 (c), and skew-sIC time courses of the study shown in a and b (d).

is a good idea to reduce the complexity of the problem using this method.

It may also be useful to simultaneously reduce the dimension of the data with whitening. The eigenvalues of the covariance matrix that are too small would then be discarded, as done in the statistical technique of principal component analysis, and often has the effect of reducing noise. Moreover, dimension reduction prevents overlearning, which is sometimes observed in ICA (22).

Clinical data. The study included 17 patients with colorectal tumor recurrences and 1 patient with liver metastasis, who were referred on the basis of clinical symptoms and radiological examinations, either CT or MRI. The final diagnosis was based on the histological data obtained from surgical specimens. None of the patients had received chemotherapy or radiation therapy at least 3 months prior to the PET study, and all patients gave their informed consent. The study was performed in accordance with the Institutional Review Board requirements.

Dynamic PET studies were performed after intravenous injection of 300-370 MBq 18F-FDG for 60 min. A 23-frame protocol was used (10x1 min, 5x2 min and 8x5 min). The 18F-FDG was prepared according to the protocol described by Toorongian *et al* (23).

A dedicated PET system (ECAT EXACT HR+; Siemens, Erlangen, Germany) was used for the patient studies. The system consists of four rings of 72 BGO detector blocks. Each block detector is divided into an 8x8 matrix, while the crystal size of an individual detector element is 4.39x4.05x30 mm^3. The system allows the simultaneous acquisition of 63 transverse slices with a theoretic slice thickness of 2.4 mm and has an axial field of view of 15.3 cm. The system was operated in two-dimensional mode (with septa extended). Transmission scans for a total of 10 min were obtained with three rotating germanium pin sources before the first radionuclide application for the attenuation correction of the acquired emission tomographic images.

All PET images were attenuation corrected, and an image matrix of 128x128 pixels was used. The images were reconstructed using an iterative reconstruction algorithm (weighted least-square method, ordered subsets, four subsets, and six iterations) and the standardized uptake values (SUV) were calculated using: SUV = tissue concentration (MBq/g)/(injected activity (MBq)/body weight (g).

The SUV calculations were done based on the last study frame (55-60 min post injection). No partial volume correction was performed, however, SUV measurements were performed on volumes of interest (VOIs) spanning several tomographic slices, instead of conventional methods that average the measured concentration over a ROI drawn in just one slice.

Results

Dynamic PET images are assumed to be the sum of components representing different structures. Each component consists of a time course and contribution ratio to the voxels (parametric image), showing the spatial distribution of the corresponding structure in the target tissue. Application of the ICA to real clinical studies aims to create a set of parametric images where different structures are readily identified. The images created by applying FastICA to the liver metastasis study are presented in Fig. 1.

A known ambiguity for ICA is that the order and variances of the independent components cannot be determined. FastICA can extract an arbitrary group of components depending on the implemented criterion and initial conditions. Contrary to the symmetric de-correlation approach, images depicting the structures of interest appear in the first independent components when using the deflation method. Moreover, they have better contrast and signal-to-noise characteristics, especially for the 'tanh' and 'gauss' non-linearity. The results obtained and initial conclusions drawn for this case have been verified by applying the same analysis to all of the colorectal tumor recurrence clinical studies.

When using the 'tanh' non-linearity, the tumor and vessels are shown in black in different images (Fig. 2) corresponding

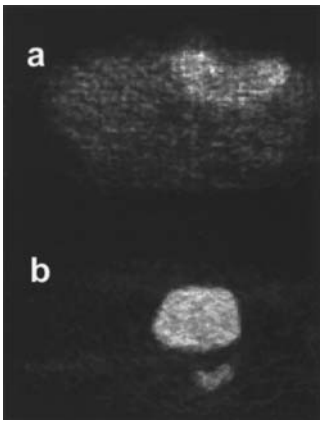

Figure 5. SUV images of the studies presented in Figs. 1 (a) and 2 (b).

to decreasing and increasing independent components, respectively (Fig. 3). In the case of 'gauss' non-linearity, tumors are presented in black and vessels in white. The corresponding ICs are both decreasing. Tumor IC always has negative values, and vessel IC is positive. Thus, the structures present in the resulting ICA images could be identified in both cases.

Fig. 4 shows the results of applying skew-sICA and skew-tICA to 2 clinical studies. The first skew-tICA image contains a bright area corresponding to the lesion, leaving the area covered by the vessels in black, whereas vessels are displayed in white and tumors in gray/black in the second skew-tICA image. However, increased noise levels complicate tumor detection in several cases (Fig. 4b). On the other hand, skew-sICA images show better noise characteristics and improve the separation of structures (Fig. 4a and c). Although no correlation has been found between the order of skew-sICA images and the structures present within them, the shape of the resulting IC time courses allows the identification of structures. In all cases, the skew-sICA images corresponding to monotonically increasing time courses display tumors in white, while monotonically decreasing time courses are related to images depicting vessels (Fig. 4d). In the case of a tumor in the vicinity of the bladder (Fig. 2), skew-sICA and skew-tICA provide comparable images, delineating the tumor slightly better than the 'tanh' and 'gauss' non-linearity FastICA approaches.

Discussion

Independent component analysis is an approach to linear decomposition and blind source separation (BSS) that is becoming increasingly popular as a tool for analyzing biomedical data (24,25). We have applied the ICA method in dynamic PET studies to reduce the initial amount of image data to a comprehensive set of parametric images, supporting visual interpretation and assisting the application of compartment modeling.

In general, all ICA and BSS algorithms have two ambiguities of arbitrary permutation and scaling, and the estimated components are ordered in a post-processing step using several criteria, e.g. according to the decreasing value of normalized kurtosis, sparseness, linear predictability, Hurst exponent, entropy, etc. However, some algorithms (e.g. AMUSE) are able to deliver online-ordered estimated components and done automatically (26,27).

Performing singular value decomposition (SVD) on the original data, as in the case of the infomax-skew-ICA, makes it possible to sort the ICs by the amount of variance of the original signal, which explain and select the components with the largest variances, thus reducing the time required for the analysis. If SVD is not applied, the deflation de-correlation approach, which estimates the ICs one-by-one, is generally more desirable than estimating all ICs in parallel (symmetric approach) when a specific IC is to be calculated (28). Using the deflation de-correlation in the FastICA technique, the computational load of the method is decreased and images with improved structure detectability are generated. Concordant with Naganawa *et al*, this approach is better suited for analyzing dynamic PET studies (7).

Among the non-linearity functions used in FastICA, 'tahn' and 'gauss' provide better quality images than 'pow3' and are able to separate structures with different kinetic characteristics such as tumors and vessels (Figs. 1 and 2). On the other hand, all structures present in the dynamic studies may be depicted in the skew-ICA images. Structures with different kinetic characteristics are easily discriminated, as they are assigned opposite values (Fig. 4c). In all cases, the identification of structures is based on the shape and sign of corresponding ICs (Figs. 3 and 4). The sign of unmixed components in ICA does not necessarily match those of the original signals, and negative ICs could therefore be generated. In all clinical data studied, ICs corresponding to structures positively displayed in the parametric images have positive values and similar shape with the time activity curves (TACs). Negative ICs correspond to negatively displayed structures with shapes symmetrically opposite to the original TACs. Although further modifications are required to exactly estimate original TACs (6,7), ICs can readily identify structures present in the images and assist visual interpretation.

The high kurtosis pdf signals extracted with standard ICA are characteristic to many natural source signals, such as speech, but there is no *a priori* reason to suppose that the pdfs of source images have high kurtosis for dynamic PET data. On the contrary, source images consisting of spatially localized features such as vascular activity surrounded by a homogeneous background have skewed pdfs (21). Incorporating physically realistic assumptions improves the robustness of ICA and has been shown to provide results for fMRI data that appear physically more plausible than those obtained with other traditional methods (5). Similarly, when skew-ICA is applied to studies with lesions in the vicinity of hot organs (e.g. bladder), the structures of interest are better discriminated (Fig. 2).

Spatial ICA separates the activity of different structures under the assumption that its distribution is spatially independent from each other. The assumption of spatial independence could be weakened by the partial-volume and spillover effects in the PET images. The time courses of different structures may also be correlated with each other because the convolution process of tracer kinetics usually relates them. In practice, the discrimination of structures is

patial than temporal ICA images, with the exception near the bladder (Fig. 2).

SUV (Fig. 5) and ICA images have comparable diagnostic performance. However, SUV offers information only on the static image frame (here at 55-60 min post-injection) under evaluation, independently of the previous image frames, whereas ICA generates images reflecting the underlying kinetics of the radiotracer uptake and offers additional time-related information. Moreover, the use of SUV as a method for classifying tissue areas as benign or malignant is still under discussion among nuclear medicine physicians and oncologists (2,3), as high metabolic activity (reflected by FDG uptake) can arise for a large number of reasons that are unrelated to tumors. A wide variety of benign disorders have high SUVs (29) including inflammation areas, a common cause of 'false positive' 18F-FDG PET scans (30). On the other hand, some malignant lesions may not necessarily have particularly high SUVs. Furthermore, high degrees of fluctuation for the SUVs from one day to another have been observed. Camera quality controls and calibrations are other parameters on which SUV measurements depend. In addition, there is currently no study on the dependence of SUVs on the type of PET camera (i.e. BGO- vs. LSO-based tomographs, etc.) or data processing prior to image formation, such as attenuation and scatter corrections, efficiency normalization methods, etc., while some studies already indicated dependence on the image reconstruction method (31). Therefore, the utility of information provided by the SUVs largely depends on its integration with all available clinical and instrumental data (32), and ICA could be an alternative to the SUV-based evaluation for studies that require dynamic PET acquisitions.

In conclusion, the results presented above showed that independent component analysis automatically generates images in which structures with different kinetic characteristics, such as tumors and vessels, are assigned opposite values and readily discriminated. Especially in cases where poor image quality due to lack of iterative image reconstruction or lesions characteristics (size, location, etc.) complicate the visual interpretation of dynamic 18F-FDG-PET datasets, ICA could improve the detectability of lesions. Moreover, it has little computational complexity and images are produced in a short time. Therefore, ICA can provide an accurate tool for the support of both the visual inspection and posterior detailed kinetic analysis of the dynamic series via compartmental or non-compartmental models by facilitating a more accurate selection of ROIs. The possibility of quantitative analysis of the dynamic PET studies using independent component analysis is under investigation.

Acknowledgements

The authors would like to thank Professor Ludwig G. Strauss and Dr Antonia Dimitrakopoulou-Strauss from the German Cancer Research Center (Heidelberg, Germany) for providing the clinical data for the experiments shown in this work and for clarifying discussions. G. Kontaxakis and A. Santos were partly supported by the EMIL Network of Excellence (European Commission, contract no. 503569).

References

1. Strauss LG and Conti PS: The applications of PET in clinical oncology. J Nucl Med 32: 623-648, 1991.
2. Keyes JW: SUV: Standard uptake or silly useless value? J Nucl Med 36: 1836-1839, 1995.
3. Huang SC: Anatomy of SUV. Nucl Med Biol 27: 643-646, 2000.
4. Strauss LG, Dimitrakopoulou-Strauss A and Haberkorn U: Shortened PET data acquisition protocol for the quantification of 18F-FDG kinetics. J Nucl Med 44: 1933-1939, 2003.
5. Stone JV, Porril J, Porter NR and Wilkinson ID: Spatiotemporal independent component analysis of event-related fMRI data using skewed probability density functions. Neuroimage 15: 407-421, 2002.
6. Lee JS, Lee DS, Ahn JY, *et al*: Blind separation of cardiac components and extraction of input function from $H_2{}^{15}O$ dynamic myocardial PET using independent component analysis. J Nucl Med 42: 938-943, 2001.
7. Naganawa M, Kimura Y, Ishii K, Oda K, Ishiwata K and Matani A: Extraction of a plasma time-activity curve from dynamic brain PET images based on independent component analysis. IEEE Trans Biomed Eng 52: 201-210, 2005.
8. Hyvärinen A and Oja E: Independent component analysis: algorithms and applications. Neural Networks 13: 411-430, 2000.
9. Huber P: Projection pursuit. Ann Stat 13: 435-475, 1985.
10. Hyvärinen A: New approximations of differential entropy for independent component analysis and projection pursuit. Adv Neural Info Process Systems 10: 273-279, 1998.
11. Hyvärinen A and Oja E: A fast fixed-point algorithm for independent component analysis. Neural Comput 9: 1483-1492, 1997.
12. Hyvärinen A: Fast and robust fixed-point algorithms for independent component analysis. IEEE Trans Neural Networks 10: 626-634, 1999.
13. Giannakopoulos X, Karhunen J and Oja E: Experimental comparison of neural ICA algorithms. Proc Int Conf Artificial Neural Networks (ICANN 98): 651-656, 1998.
14. Vigário R, Jousmäki V, Hämäläinen M, Hari R, and Oja E: Independent component analysis for identification of artifacts in magnetoencephalographic recordings. Adv Neural Info Process Systems 10: 229-235, 1998.
15. Kiviluoto K and Oja E: Independent component analysis for parallel financial time series. Proc Int Conf Neural Info Process (ICONIP 98) 2: 895-898, 1998.
16. Hyvärinen A: Sparse code shrinkage: denoising of non-gaussian data by maximum likelihood estimation. Neural Comput 11: 1739-1768, 1999.
17. Ristaniemi T and Joutsensalo J: On the performance of blind source separation in CDMA downlink. Proc Int Workshop Independent Component Anal Signal Separation (ICA 99): 437-441, 1999.
18. Bell A and Sejnowski T: An information-maximization approach to blind separation and blind deconvolution. Neural Comput 7: 1129-1159, 1995.
19. Nadal JP and Parga N: Non-linear neurons in the low noise limit: a factorial code maximizes information transfer. Network 5: 565-581, 1994.
20. Williams P: A Marquardt algorithm for choosing the step-size in backpropagation learning with conjugate gradients. Cognitive Science Research Paper CSRP 229. University of Sussex, 1991.
21. Suzuki K, Kiryu T and Nakada T: Fast and precise independent component analysis for high field fMRI time series tailored using prior information on spatiotemporal structure. Human Brain Mapping 15: 54-66, 2002.
22. Hyvärinen A, Särelä J and Vigário R: Spikes and bumps: Artefacts generated by independent component analysis with insufficient sample size. Proc Int Workshop Independent Component Anal Signal Separation (ICA 99): 425-429, 1999.
23. Toorongian SA, Mulholland GK, Jewett DM, Bachelor MA, Kilbourn MR: Routine production of 2-deoxy-2(F-18)fluoro-D-glucose by direct nucleophilic exchange on a quaternary 4-amino-pyridinium resin. Nucl Med Biol 3: 273-279, 1990.
24. Porril J, Stone J, Berwick J, Mayhew J and Coffey P: Analysis of optical imaging data using weak models and ICA. In: Advances in Independent Component Analysis. Girolami M (ed). Springer-Verlag, pp217-233, 2000.
25. McKeown M: Detection of consistently task-related activations in fMRI data with hybrid independent components. Neuroimage 11: 24-35, 2000.

26. Tong L, Liu RW, Soon VC and Huang YF: Indeterminacy and identifiability of blind identification. IEEE Trans Circuits Systems 38: 499-509, 1991.
27. Molgedey L and Schuster HG: Separation of a mixture of independent signals using time delayed correlations. Phys Rev Lett 72: 3634-3636, 1994.
28. Hyvärinen A, Karhunen J and Oja E: ICA by maximization of nongaussianity. In Independent Component Analysis. Wiley Interscience, New York, pp165-202, 2001.
29. Lapela M, Eigtved A, Jyrkkio S, *et al*: Experience in qualitative and quantitative FDG PET in follow-up of patients with suspected recurrence from head and neck cancer. Eur J Cancer 36: 858-867, 2000.
30. Strauss LG: F-18 deoxyglucose and false-positive results: a major problem in the diagnostics of oncological patients. Eur J Nucl Med 23: 1409-1415, 1996.
31. Visvikis D, Cheze-LeRest C, Costa DC, Bomanji J, Gacinovic S and Ell PJ: Influence of OSEM and segmented attenuation correction in the calculation of standardised uptake values for [18F]FDG PET. Eur J Nucl Med 28: 1326-1335, 2001.
32. Lucignani G, Paganelli G and Bombardieri E: The use of standardized uptake values for assessing FDG uptake with PET in oncology: a clinical perspective. Nucl Med Commun 25: 651-656, 2004.