Use of Excess Height and Cluster Extent in Subtraction SPECT

Kristof Baete, Johan Nuyts, *Member*, *IEEE*, Wim Van Paesschen, Alex Maes, Shivani Ghoorun, Paul Suetens, *Member*, *IEEE*, and Patrick Dupont

Abstract— **Subtraction of ictal and interictal SPECT perfusion images of the brain has the potential of locating the epileptogenic region. This region generally shows large differences between both images. However, differences can also be induced by noise in the projection data. We hypothesized that the extent, besides the intensity, of observed clusters of voxels in thresholded subtraction images, is an important parameter in the classification of clusters into real perfusion differences and noise induced differences. To test this hypothesis, we performed a number of simulation experiments. Using a Monte Carlo approach, we constructed cumulative distribution functions (CDF) of the excess height (i.e. the largest difference in a cluster) and the cluster extent under the condition of no perfusion change (i.e. only noise induced clusters). The reproducibility of the CDF curves was shown using measured patient data. Furthermore, a 3D brain software phantom experiment was used to examine the detection and classification of an induced region of hyperperfusion. In a first experiment, we compared two detection criteria: detection of the induced hyperperfusion based on the observed cluster with the largest excess height and based on the observed cluster with the largest extent. Detection based on the largest extent showed a better sensitivity. In a second experiment, we assigned to every observed cluster a probability, derived from the CDF curves, for excess height and extent. For different probability thresholds, sensitivity and specificity of the detection of the induced hyperperfusion based on its probability for excess height and cluster extent was measured. These measurements were combined in receiver operating characteristic (ROC) curves. These ROC curves showed a better performance when using classification based on cluster extent. We conclude that the cluster extent is an important parameter in the characterization of clusters in thresholded subtraction of perfusion SPECT images of the brain.**

Keywords—**Single photon emission computed tomography, image processing, epilepsy**

I. Introduction

Single photon emission computed tomography (SPECT) of the cerebral perfusion in patients with epilepsy is useful in the presurgical evaluation [1–5]. In fact, it is known that during a seizure the perfusion in the epileptogenic region is increased and between seizures it is decreased in about half of the cases [6]. Examination of perfusion images acquired during (ictal) and between (interictal) seizures can then lead to the detection of the epileptogenic region(s). However, in that way some subtle but significant differences may not be appreciated visually and it has been shown that localization can be improved when these images are analyzed on a quantitative basis [7]. Image post-processing techniques and quantitative analysis can thereby enhance the detection of real differences between two functional images of the brain.

The method of ictal-interictal subtraction is a straightforward approach that was proposed to partially solve the detection problem of differences in functional brain images [8]. Subtraction SPECT co-registered to MRI (SISCOM) even increases the diagnostic yield and allows functional and anatomical correlation [1–3, 9]. In this technique, the subtraction image is thresholded and only the largest differences are visualized. Voxels exceeding the threshold are then likely to belong to regions with epileptogenic or normal physiological changes. Unfortunately, noise in the measured projection data can eventually induce clusters of supra threshold voxels in the subtraction image as well. This noise in the data is caused by the decay of the limited amount of radio-pharmaceutical that can be administered to a patient and the limited scanning duration.

The use of a quantitative post-processing method could improve the diagnostic value of subtraction SPECT in clinical practice by making a probabilistic discrimination between clusters induced by noise on the one hand and clusters induced by physiological or pathological processes on the other hand. This has been the aim of many groups working in the field of statistical parametric mapping (SPM) analysis of functional brain images. This technique demands statistical modeling of the problem and, although it is an excellent and powerful technique, it has some disadvantages for this type of problem [10, 11]. Here, the number of scans per study is much too small which results in conservative inferences and an increased amount of false negative predictions. Moreover, when the statistical image is constructed, there is no incorporation of image processing steps, i.e. acquisition, reconstruction and post-processing. Altering one of these could generate subtraction images with other characteristics, thereby influencing the significance of the observed differences.

With the disadvantages of the theoretical model in mind, supra threshold cluster tests have been explored before, using a simulation approach [12, 13]. In those studies, experiments attempted to mimic real null statistic images using

Manuscript received October 23, 2001; revised May 7, 2002. This work was supported by grants IDO-99/005 and BIL-99/41.

K. Baete, J. Nuyts, A. Maes and P. Dupont are with the Department of Nuclear Medicine, University Hospital Gasthuisberg, B-3000 Leuven, Belgium (e-mail: kristof.baete@uz.kuleuven.ac.be). P. Dupont is post-doctoral researcher of the Flemish Fund for Scientific Research (FWO).

W. Van Paesschen is with the Department of Neurology, University Hospital Gasthuisberg, B-3000 Leuven, Belgium.

S. Ghoorun is with the Department of Nuclear Medicine, Tygerberg Hospital, Cape Town, South Africa.

P. Suetens is with the Laboratory for Medical Image Computing, Radiology-ESAT/PSI, University Hospital Gasthuisberg, B-3000 Leuven, Belgium.

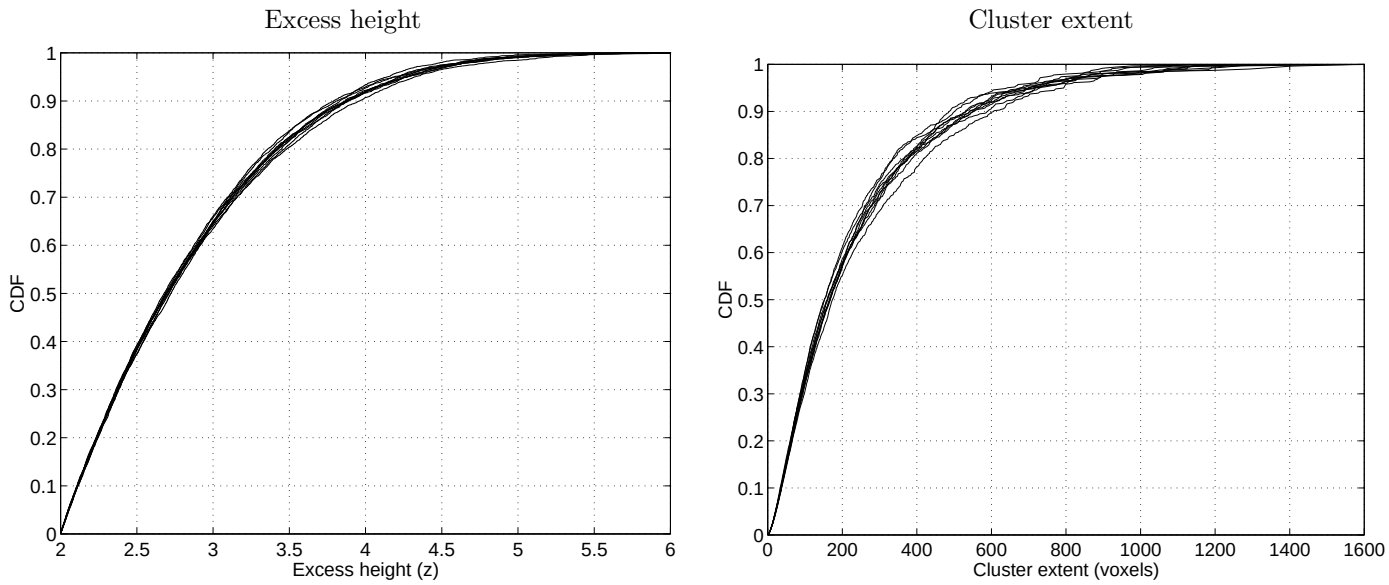

Fig. 1. Each curve represents the empirical cumulative distribution function (CDF) of the excess height (*left*) and cluster extent (*right*) of clusters found in 150 thresholded subtraction images. These images are subtractions of reconstructed noise realizations of the projection data of an individual. Ten individuals were included in this study.

smoothed white noise Gaussian fields with the variance of noise and the full width at half maximum (FWHM) of the smoothing kernel chosen to match that of real difference images. However, there was no incorporation of acquisition parameters, reconstruction method or characteristic features of further image processing.

In this study, based on a Monte Carlo approach, we propose an empirical method for the assignment of a probability for excess height and for extent to every observed cluster of voxels in a thresholded subtraction image under the null hypothesis of no real perfusion change. The excess height of a cluster is defined as the largest difference shown in the voxels belonging to that cluster and extent is expressed as the number of voxels in a cluster. The construction of the method is based on our standard epilepsy imaging protocol but it can be used for other protocols as well. This makes the use of the method only valid for data processed with the same protocol. However, the advantage of our approach lies in the fact that characteristics of acquisition, reconstruction and image processing are incorporated in the method.

We assessed reproducibility of our method by means of simulation, based on Poisson noise realizations of measured projection data. We assumed that construction of noise realizations is the same as doing repeated measurements without physiological or physical changes. Thereby, we neglected the possible contribution of misregistration errors. However, registration algorithms perform well in practice [14–16].

Subsequently, this empirical method was used for the investigation of our main hypothesis, i.e. the cluster extent is an important parameter, besides excess height, for the characterization of significant differences. We tested this hypothesis using a 3D software phantom simulation in a number of experiments.

II. Materials and methods

A. Distribution of cluster extent and excess height

We determined the distribution of the measured excess height and extent of observed clusters in a number of thresholded subtraction images.

First, a trained nuclear medicine physician examined the reconstruction images of patients who received a brain-SPECT using the radio-pharmaceutical ^{99m}Tc-ethyl cysteinate dimer (ECD) and selected 10 patients who showed a normal perfusion pattern. Data were acquired using a triple headed gamma-camera (Triad XLT, Trionix) equipped with low-energy, ultra-high resolution parallel-hole collimators, in a 128 × 64 matrix over 360 degrees, with 120 views obtained at 3 degrees intervals for 15 seconds per view, using a non-circular orbit. The pixel size measured 2.54 mm and the energy setting was 140 keV with a 20 % window.

Thereafter, the projection data of every individual were used to make 300 pseudo-random Poisson noise realizations. Each noise realization was then reconstructed according to the standard clinical procedure, i.e. by prefiltering the projections with a 2D Butterworth filter (cut-off frequency = 0.154 cyc/pix, power = 10) followed by a filtered backprojection using the ramp filter. Attenuation correction was performed using Chang's method [17] with $\mu = 0.12\,\mathrm{cm}^{-1}$. For each individual, the reconstruction images were randomly divided into 150 pairs. Furthermore, a brain mask was defined to identify cerebral voxels for further analysis.

The reconstructions were normalized by the mean cerebral voxel intensity to account for differences in e.g. the injected dose and radioactive decay. Subtraction for each pair of images was done only for voxels inside the brain mask of the individual. The difference image was convolved

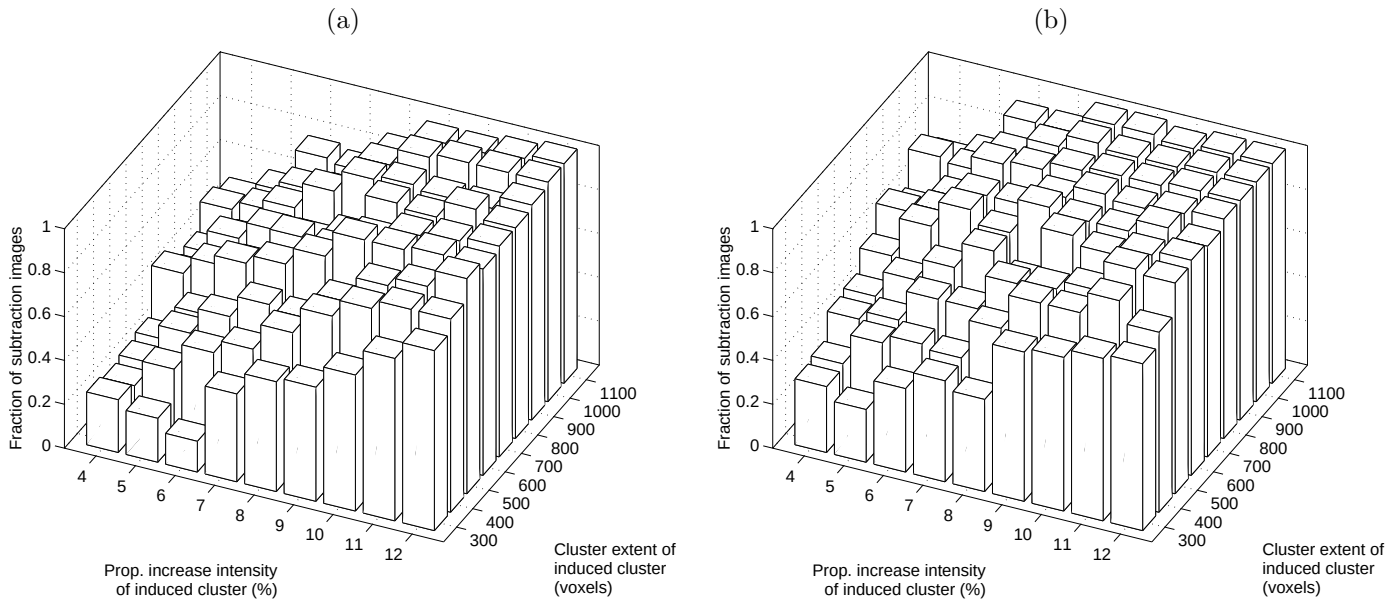

Fig. 2. Each bar represents the fraction of subtraction images in which the true retrieved cluster (TRC) was the observed cluster with the highest excess (a), respectively the largest extent (b).

with a Gaussian smooth kernel with a FWHM of 5 pixels. The mean μ and the standard deviation σ of the differences Δ between each pair were calculated. The differences were transformed into a standard z-score using $z = (\Delta - \mu)/\sigma$. Supra threshold images were obtained by thresholding at $z = 2$. Supra threshold clusters were defined using an 18-connectivity scheme [18]. The size or extent of every cluster inside the brain was given by the number of voxels. Also, the excess height of the observed clusters was measured.

For each individual, a cumulative distribution function (CDF) was computed for the measured values of cluster extent and excess height of supra threshold clusters in the 150 subtraction images. Statistical analysis was applied onto the curves to assess inter-individual differences. Finally, data of all ten individuals for each parameter were combined and a general mean CDF was computed.

B. Phantom simulation experiments

We studied the performance of the parameters excess height and cluster extent in the detection of induced clusters with two software phantom simulation experiments.

The average of four 3D-MPRAGE T1-weighted magnetic resonance images of one individual was segmented using SPM99 (http://www.fil.ion.ucl.ac.uk/spm). The gray matter segmentation map was brought to SPECT resolution by image registration [14] with a brain ^{99m}Tc-ECD SPECT of the same individual. The same SPECT acquisition parameters and image dimensions were used as described above. The binary segmentation map was then uniformly scaled in intensity so that a clinical realistic count level was reached when projecting the image with the incorporation of attenuation. This image was used as the baseline phantom.

The activation phantom was made by adding a cluster, acting as a hyperperfusion, to the baseline phantom. Therefore, in the temporal pole of the segmentation map, a region frequently found to be involved in epilepsy, we fixed the central point of a sphere with adjustable radius. Voxels common to the gray matter segmentation image and the sphere were then proportionally increased in intensity. The induced cluster was always restricted to the gray matter and deviated from a spherical shape when larger than the gray matter boundary. Thereby, the activation phantom was characterized by the effective size in voxels and the proportional increase of intensity of the induced cluster. The effective cluster size used in this experiment was from 300 up to 1100 voxels in steps of 100 voxels. The proportional increase of intensity was from 4 to 12 percent in steps of 1 percent. We will call this the sample space.

The acquisition process was simulated by projection of the baseline and activation phantom using the same uniform attenuation mask. For every combination of proportional increased intensity and size of the induced cluster, 50 pseudo-random Poisson noise realizations were computed for the projection data of both phantoms. Noise realizations of the projections were reconstructed with the standard clinical procedure, as described above, and every pair was normalized, subtracted and thresholded at $z = 2$. Thereafter, each subtraction image was analyzed regarding the number of clusters, their extent, excess height and location in the brain. In every thresholded subtraction image, the cluster (if existing) corresponding to the induced cluster in the activation phantom was identified. This cluster was named the true retrieved cluster (TRC).

Extreme differences in size or height are easily perceptible. In a first experiment, we measured the sensitivity of detecting the induced cluster in the subtraction image of a baseline and activation phantom, when looking at either the supra threshold cluster with the largest extent or the highest intensity. For every induced cluster in the activation phantom, with a given size and proportional increase

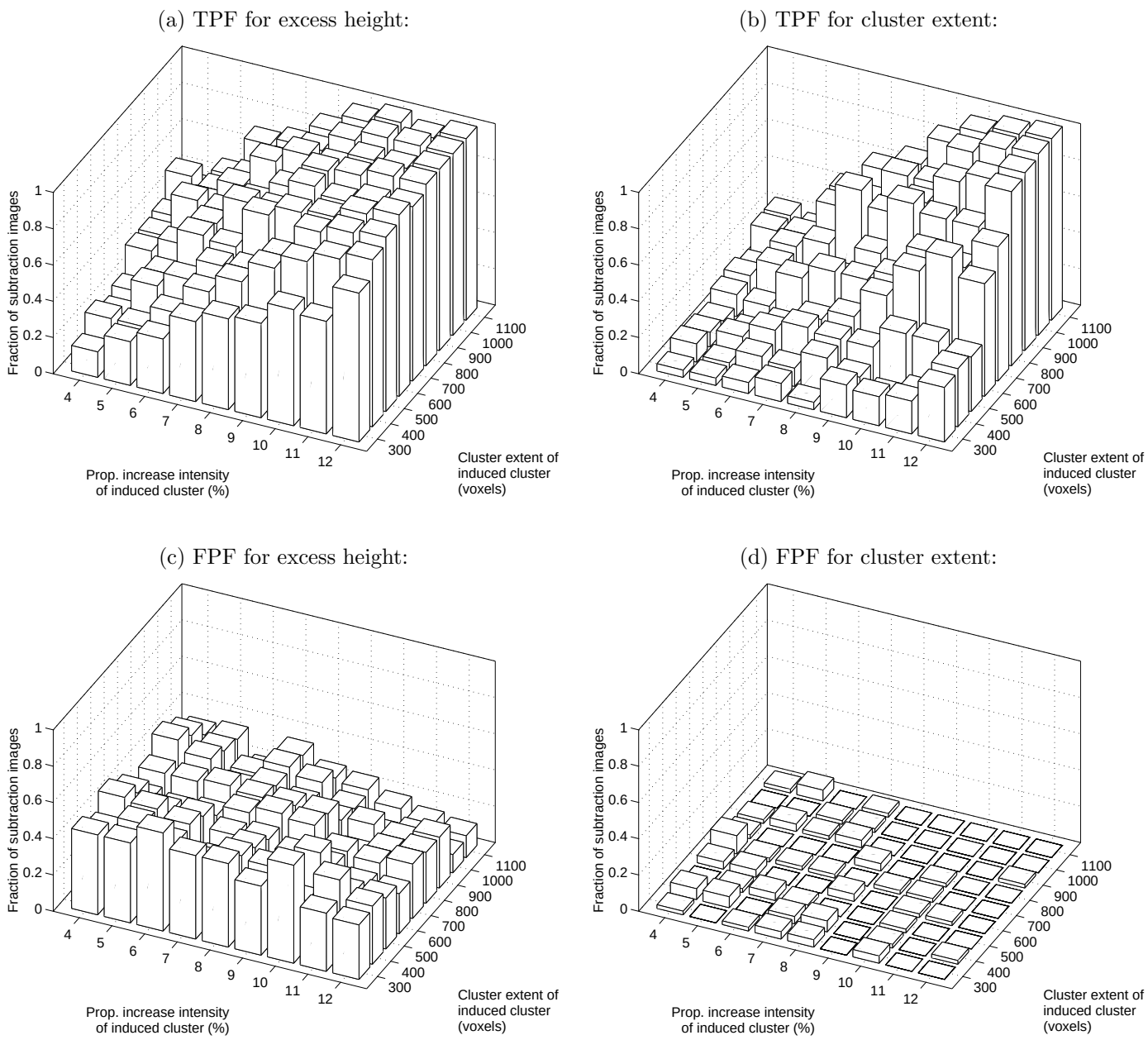

Fig. 3. For different samples in size and height of the induced cluster, the true positive fraction (TPF) for excess height (a) and cluster extent (b), and the false positive fraction (FPF) for excess height (c) and cluster extent (d) is calculated. The critical probability threshold p_c was set to 0.95.

of intensity, the fraction of subtraction images over 50 noise realizations that satisfied this criterion was measured. In that way, we quantified the detectability of an induced cluster based on extreme values of the measurable parameters excess height or cluster extent.

In a second experiment, we analyzed the specificity and sensitivity of both parameters in the classification of clusters. For that purpose, we used the simulated distribution of measured parameter values in subtraction images under noise conditions and without any activation. CDF curves for the excess height and extent of observed clusters in the phantom simulation experiment were computed using the method described in the beginning of this section. Therefore, 550 pairs of noise realizations of the projection data of the baseline phantom were analyzed.

Subsequently, the CDF curves were used to assign a probability for the measured excess height and cluster extent to each observed cluster in a thresholded subtraction image of the phantom simulation experiment. In fact, $\mathrm{CDF}(x) = \mathrm{Prob}(X \leq x)$, for the random variable X which is either the excess height or the extent of clusters under the condition of no real perfusion change (i.e. only noise). Then, $\mathrm{CDF}(x)$ is the chance of observing a cluster in a thresholded subtraction image with a value less than x for either one of the parameters under the condition of only noise. In the phantom simulation experiment, we chose different critical probability thresholds $0 \ll p_c < 1$. Then, clusters with a probability for excess height, respectively

cluster extent, above p_c were considered significant for the parameter excess height, respectively cluster extent. For several values of p_c and for each parameter, the specificity and sensitivity of the detection of the induced hyperperfusion was calculated.

The fraction of subtraction images, where the true retrieved cluster (TRC) was considered significant, was computed. This represents the true positive fraction (TPF) or sensitivity for the p_c of the used detection parameter. The fraction of subtraction images in which a cluster, except the TRC, was found to be significant was also computed. This represents the false positive fraction (FPF) or one minus the specificity for the p_c of the used detection parameter. These two values were calculated for different sizes and heights of the induced cluster and for different probability thresholds. For each parameter, the mean TPF was computed over all applied sizes and excess heights (i.e. all samples in the sample space) of the induced cluster. Similarly, the mean FPF was computed. To compare the performance of the parameters cluster extent and excess height, a ROC curve was constructed based on the mean TPF and mean FPF for different probability thresholds.

III. Results

A. Cumulative distribution functions

The cumulative distribution curves of the measured excess height and extent of observed clusters in thresholded subtraction images of 10 individuals are shown in Fig. 1. These curves show that noise in the projection data can induce considerably large differences in size and height in a subtraction image. A Kolmogorov-Smirnov test was performed to assess inter-individual robustness, using the data sets of the individuals with the most remote CDF curves seen in Fig. 1. The test showed $p > 0.10$ for both excess height and cluster extent. This means that the distributions of both parameters are not significantly different between individuals.

B. Phantom simulation experiment

We analyzed 81 samples with different size and proportional increase of intensity of the induced cluster. For each sample, the fraction of subtraction images was computed in which the true retrieved cluster was the cluster with the highest intensity respectively the largest extent. Results are shown in Fig. 2.

In the second phantom simulation experiment, firstly, the critical probability threshold p_c was set to 0.95 and the true and false positive fraction of subtraction images for all samples was computed based on excess height and cluster extent. The results are shown in Fig. 3. The mean values of the TPF and FPF over all samples in the sample space are:

($p_c = 0.95$)	Excess height	Cluster extent
mean TPF	0.70	0.40
mean FPF	0.33	0.02

Secondly, the probability threshold p_c was varied between 0.1000 and 0.9980 and the mean TPF and mean FPF of all subtraction images was computed. The ROC-curve containing those results is shown in Fig. 4. The area under the ROC-curve, based on the excess height, is 0.74. Based on the cluster extent, the area is 0.83. Statistical comparison of the area under the ROC-curves, using the ROC-analysis software package ROCKIT 0.9 (http://www-radiology.uchicago.edu/krl) [19], showed significant difference ($p < 0.0001$).

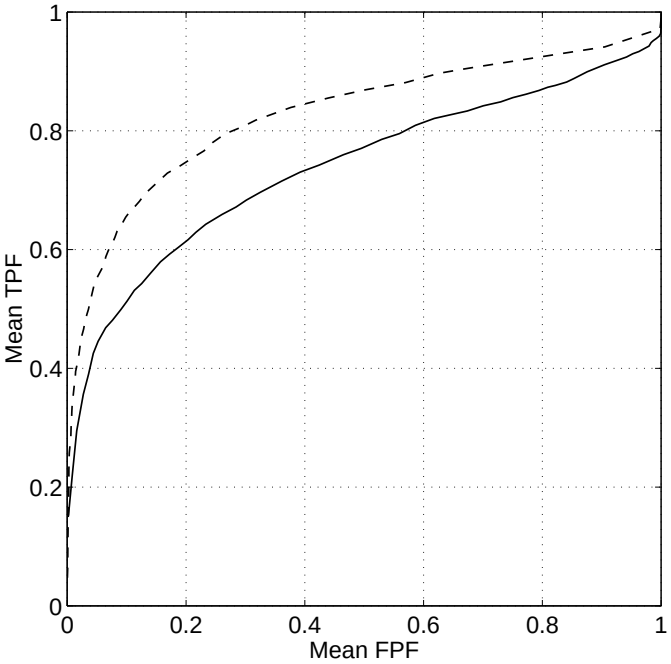

Fig. 4. ROC-curves based on different probability thresholds p_c for the parameters excess height (*solid line*) and cluster extent (*dashed line*). Results are determined on the mean TPF and FPF values over all samples.

IV. Discussion

We have described the construction of empirical CDF curves for the parameters excess height and cluster extent of observed clusters in thresholded subtraction images (Fig. 1). Conceptually, this Monte Carlo approach and construction of these curves is straightforward. Unlike previous studies [12,13], our simulations were based on measured projection data. Even though there is a known considerable inter-individual variation between brain perfusion SPECT images, the method proved to be robust for a data set of 10 normal individuals. In practice, this Monte Carlo approach requires substantial computer time. Mainly for that reason few empirical approaches have been attempted before [20]. However, computation of the CDF curves has to be done only once, but depends on the acquisition and reconstruction procedure. Whenever the standard procedure (acquisition or reconstruction) would change, new CDF curves have to be computed.

The first experiment of the phantom simulation study investigated the detection performance of the parameters excess height and extent using a specific criterion, viz. the largest parameter values. It showed (see Fig. 2) that a bet-

ter detection of the induced cluster could be achieved when looking at the largest cluster rather than at the cluster with the greatest excess height in the thresholded subtraction image. The mean sensitivity over the sample space based on extent is about 11 % higher compared to the mean value based on excess height. For less intense induced clusters, the increase of mean sensitivity based on extent can even reach 25 % higher compared to excess height. This proves what has been assumed before, viz. close congregation of voxels have more significance than high intensity values [9].

In the second experiment of the phantom simulation, we chose initially a 0.95 probability threshold for excess height and cluster extent. Inferences based on this significance level are commonly used in statistical analysis. The computed TPF and FPF of subtraction images for different samples in the sample space at this significance level are shown in Fig. 3(a)–(d). Striking is the FPF bar chart of excess height, which shows an overall high amount of false positives at the 0.95 significance level, compared to the FPF bar chart of cluster extent. The difference between both bar charts could be caused by different behavior of the detection parameter at the chosen significance level. Therefore, the critical probability threshold p_c was varied to assess the difference in overall performance between excess height and extent. We chose to construct the ROC curves based on the mean TPF and mean FPF over the sample space. The ROC curve (Fig. 4) of extent showed a much better characterization of clusters and was proven to be significantly different from that of excess height. This shows that the accuracy of extent in the detection of an induced hyperperfusion in a thresholded subtraction image is better than that of excess height. The area under the ROC curve of the cluster extent is about 12 % higher compared to that of the excess height. Besides the analysis over the whole sample space, we divided the sample space in subregions and computed ROC curves using the mean TPF and mean FPF over the subregion. For every subregion, the area under the ROC curve for the cluster extent was always greater than that for the excess height. Based on a visual assessment study, others [5] found similar results concerning sensitivity and specificity of subtraction SPECT when cluster extent was used as a criterion.

One could question the choice of the sample space we made in this experiment. However, the sample space was chosen in such a way that detection of the induced cluster was visually not obvious. Detection of induced hyperperfusions with larger values for proportional increase and size seemed to be of no problem. This has been examined visually with several activation phantoms and their corresponding subtraction images.

One should consider that large differences between, e.g. ictal and interictal images, can influence the histogram of the subtraction image and hence the generation of supra threshold clusters. This will contribute to a lower FPF of the remaining observed clusters. However, in such cases these differences are obvious in a visual analysis. In the case of more moderate differences this effect is less prominent.

Further analysis of our method in clinical practice is required to evaluate its value in the characterization of supra threshold clusters in subtraction SPECT images.

V. Conclusion

We showed that cluster extent is an important parameter in the characterization of clusters in a thresholded subtraction of perfusion SPECT images of the brain.

VI. Acknowledgments

We wish to thank Dirk Vandermeulen, Siddharth Srivastava and Dirk Bequé for useful discussions regarding this study.

References

[1] T. J. O'Brien, E. L. So, B. P. Mullan, M. F. Hauser, B. H. Brinkmann, N. I. Bohnen, D. Hanson, G. D. Cascino, C. R. Jack, and F. W. Sharbrough, "Subtraction ictal SPECT co-registered to MRI improves clinical usefulness of SPECT in localizing the surgical seizure focus," *Neurology*, vol. 50, pp. 445–454, 1998.

[2] T. J. O'Brien, E. L. So, B. P. Mullan, M. F. Hauser, B. H. Brinkmann, C. R. Jack, G. D. Cascino, F. B. Meyer, and F. W. Sharbrough, "Subtraction SPECT co-registered to MRI improves postictal SPECT localization of seizure foci," *Neurology*, vol. 52, pp. 137–146, 1999.

[3] T. J. O'Brien, E. L. So, B. P. Mullan, G. D. Cascino, M. F. Hauser, B. H. Brinkmann, F. W. Sharbrough, and F. B. Meyer, "Subtraction peri-ictal SPECT is predictive of extratemporal epilepsy surgery outcome," *Neurology*, vol. 55, pp. 1668–1677, 2000.

[4] P. Véra, A. Kaminska, C. Cieuta, A. Hollo, J. L. Stiévenart, I. Gardin, D. Ville, J. F. Mangin, P. Plouin, O. Dulac, and C. Chiron, "Use of subtraction ictal SPECT co-registered to MRI for optimizing the localization of seizure foci in children," *J. Nucl. Med.*, vol. 40, pp. 786–792, 1999.

[5] M. V. Spanaki, S. S. Spencer, M. Corsi, J. MacMullan, J. Seibyl, and I. G. Zubal, "Sensitivity and specificity of quantitative difference SPECT analysis in seizure localization," *J. Nucl. Med.*, vol. 40, pp. 730–736, 1999.

[6] S. F. Berkovic and M. R. Newton, "Single photon emission computed tomography," in *Epilepsy: a comprehensive textbook*, J. Engel and T.A. Pedley, Lippincott-Raven Publ., Philadephia, 1997, pp. 969–975.

[7] M. D. Devous, R. A. Thisted, G. F. Morgan, R. F. Leroy, and C. C. Rowe, "SPECT brain imaging in epilepsy: a meta-analysis," *J. Nucl. Med.*, vol. 39, pp. 285–293, 1998.

[8] I. G. Zubal, S. S. Spencer, K. Imam, J. Seibyl, E. O. Smith, G. Wisniewski, and P. B. Hoffer, "Difference images calculated from ictal and interictal technetium-99m-HMPAO SPECT scans of epilepsy," *J. Nucl. Med.*, vol. 36, pp. 684–689, 1995.

[9] T. J. O'Brien, M. K. O'Connor, B. P. Mullan, B. H. Brinkmann, D. Hanson, C. R. Jack, and E. L. So, "Subtraction ictal SPET co-registered to MRI in partial epilepsy: description and technical validation of the method with phantom and patient studies," *Nuc. Med. Commun.*, vol. 19, pp. 31–45, 1998.

[10] P. D. Acton and K. J. Friston, "Statistical parametric mapping in functional neuroimaging: beyond PET and fMRI activation studies," *Eur. J. Nucl. Med.*, vol. 25, pp. 663–667, 1998.

[11] B. H. Brinkmann, T. J. O'Brien, D. B. Webster, B. P. Mullan, P. D. Robins, and R. A. Robb, "Voxel significance mapping using local image variances in subtraction ictal SPET," *Nuc. Med. Commun.*, vol. 21, pp. 545–551, 2000.

[12] P. E. Roland, B. Levin, R. Kawashima, and S. Åkerman, "Three-dimensional analysis of clustered voxels in ^{15}O-butanol brain activation images," *Hum. Brain Map.*, vol. 1, pp. 3–19, 1993.

[13] J.-B. Poline and B. M. Mazoyer, "Analysis of individual positron emission tomography activation maps by detection of high signal-to-noise-ratio pixel clusters," *J. Cereb. Blood Flow Metab.*, vol. 13, pp. 425–437, 1993.

[14] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Trans. Med. Imag.*, vol.16, no. 2, pp. 187–198, 1997.

[15] D. L. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes, "Medical image registration," *Phys. Med. Bio.*, vol. 46, pp. R1–R45, 2001.

[16] L. Thurfjell, Y. H. Lau, J. L. Andersson, and B. F. Hutton, "Improved efficiency for MRI-SPET registration based on mutual information," *Eur. J. Nucl. Med.*, vol. 27, pp. 847–856, 2000.

[17] L. T. Chang, "A method for attenuation correction in radionuclide computed tomography," *IEEE Trans. Nucl. Sci.*, vol. 25, pp. 638–643, 1978.

[18] K. J. Worsley, A. C. Evans, S. Marrett, and P. Neelin, "A three-dimensional statistical analysis for CBF activation studies in human brain," *J. Cereb. Blood Flow Metab.*, vol. 12, pp. 900–918, 1992.

[19] C. E. Metz, B. A. Herman, and C. A. Roe, "Statistical comparison of two ROC-curve estimates obtained from partially-paired datasets," *Med. Decis. Making*, vol. 18, pp. 110–121, 1998.

[20] A. P. Holmes, "Statistical issues in functional brain mapping," *Ph.D. Thesis*, pp. 123–125, 1994.