Choquet integrals as projection operators for quantified tomographic reconstruction

Agnès Rico[a,*], Olivier Strauss[b], Denis Mariano-Goulart[c]

[a]*LIRIS Equipe MA2D, Université Claude Bernard Lyon I, 43 bld du 11 novembre 1918, 69622 Villeurbanne, France*
[b]*LIRMM Université Montpellier II, 61 rue Ada, 34392 Montpellier cedex 5, France*
[c]*CHU LAPEYRONIE, Service Central de Médecine Nucléaire,371, avenue du doyen Gaston Giraud 34295 Montpellier cedex 5, France*

Abstract

In this paper, we propose to investigate and analyze a new method for performing quantified projection and back-projection in emission tomography. This method is based on using non-summative kernels, capacities and asymmetric Choquet integral to obtain imprecise projected values (i.e. intervals instead of usual reconstructed pixel values).

Validation studies using numerical and physical single photon computed emission tomography (SPECT) phantoms were used to demonstrate links between the length of these reconstructed intervals and the stochastic noise level in reconstructed slices.
© 2008 Elsevier B.V. All rights reserved.

Contents

* Corresponding author. Tel.: +33 4 72 43 16 56; fax: +33 4 72 43 10 44.

E-mail addresses: agnes.rico@univ-lyon1.fr (A. Rico), strauss@lirmm.fr (O. Strauss), d-mariano_goulart@chu-montpellier.fr (D. Mariano-Goulart).

1. Introduction

Diagnosis, prognostic assessment and therapeutic monitoring of various cancers, cardiovascular or neurological diseases require in-vivo evaluation of some biological or metabolic functions in patients. The recent development of specific radioactive tracers in single photon computed emission tomography (SPECT) or in positron emission tomography (PET) have made such measurements possible through tomographic reconstruction of projections acquired by these imagers. The reconstruction algorithms, along with specific algorithms dedicated to the correction of attenuation artefacts or imager responses, which are available in routine clinical settings, provide increasingly accurate reconstructions.

However, whatever the algorithm used, the ill-posedness of the reconstruction problem leads to unstable solutions and the various regularisation functions suggested in the literature are difficult to adjust in routine clinical settings. One of the practical challenges, in nuclear tomographic image analysis, is to be able to robustly estimate image reconstruction error induced by discretization and measurement variations due to acquisition Poisson noise. Robust error estimation enables the physician to objectively compare reconstructed activities within two different regions of interest in a reconstructed slice. The reconstruction error could be estimated by repeating the reconstruction process using a large number of acquisitions. However, experimental constraints limit the number of experiments, and such a method does not apply in clinical settings. Usual reconstruction methods try to transfer knowledge on the nature of the noise acquired in projections to the reconstructed voxels. Many of these reconstruction noise estimation methods usually assume that the reconstructed image is close enough to the real image to use a Jacobian-based variance transfer [22,18,2]. For this method to be efficient, the residual random error must have suitable statistical properties (i.e. zero-mean, close to normal distribution, etc.) which is somewhat difficult.

We are currently investigating a completely different approach involving a new formulation of the tomographic reconstruction problem. We aim at designing reconstruction algorithms in which error estimation will be part of the estimation process. This paper presents one of the very first steps of this work, i.e. a new projection operator, based on non-summative kernels and non-additive confidence measures. The Choquet integral [3], which is a key tool for aggregating values associated with non-additive confidence measure, is used as projection operator. Using a concave capacity and its conjugate, we obtain an interval of projected value instead of a single value [10]. We stress that quantitative information associated with the length of this interval value is highly correlated with the level of the (back-) projected noise. Then we come up with a method for quantifying the noise level that is simpler than the conventional method using Jacobian-based variance transfer. The concept underlying this work is as follows: when estimating a piece of information from a set of noisy data with a kernel-based approach, variations due to measurement error can be estimated by slightly varying the shape (but not the size) of the kernel. Such a variation can be simply achieved by replacing a conventional summative kernel by a non-summative kernel approach [13].

The paper is organized as follows: in Section 2, we briefly review the methods for computing operators in the additive case. In Section 3, we present the framework and notations: the first subsection deals with summative versus non-summative kernels. The second one presents the Choquet integral according to a capacity and to its conjugate capacity and their relations. The third one presents the asymmetric Choquet integral as an aggregation operator. Section 4 is dedicated to the presentation of the new projection and backprojection operators and their properties. Section 5 describes how the projection operator can be implemented. Using simulated and physical SPECT phantoms, in Section 6 we show the close correlation between the length of the intervals given by this alternative method and the random noise in the projections. Even though the results presented in this paper were obtained using emission tomographic data, the theoretical framework presented here has a wide range of applications and may be used for various tomographic devices.

2. Kernel-based discrete tomography

Since the measured activity in a particular direction (ρ, θ) consists of a stochastic sum of the activity along (or around) a line, the Radon transform is usually used to model the physical phenomenon [15]. Let $I(x, y)$ be the slice to be reconstructed, and $S(\rho, \theta)$ be the integral of the stochastic activity $I(x, y)$ along a line whose polar coordinate

is (ρ, θ). The analytical expression of the Radon transform is

$$S(\rho, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(x, y)\delta(x \cos \theta + y \sin \theta - \rho)\, \mathrm{d}x\, \mathrm{d}y,$$

where δ is the Dirac delta distribution. The activity $S(\rho, \theta)$, that has been detected, is the signature of a sought after density (Hounsfield number, radio-activity, etc.). However, the analytical expression of the Radon transform has to be modified to account for measurement noise, missing data, discretization and quantification. Moreover, the problem of practical interest is to find an implementation of a discrete reconstruction algorithm.

The so-called discrete Radon transform [11,12] makes use of summative kernels [1,24] for handling the interplay between continuous and discrete. A summative (or probabilistic) kernel of $\mathbb{R}^m$ is an integrable function κ that satisfies: $\int_{w \in \mathbb{R}^m} \kappa(w)\, \mathrm{d}w$. It is usually separable, centered, symmetric, bounded and positive. A summative kernel κ is usually scaled by a bandwidth parameter that determines the local amount of smoothing [24]. The role of the kernel is to define a weighted neighborhood of each sampled location, thus providing interplay between continuous and discrete. But it also makes use of a hypothesized ergodicity to reduce the effect of random noise by locally regularizing the reconstruction process.

Let I_n be the sought after value of the nth pixel centered on the location (x_n, y_n) of the digital image to be reconstructed related by the kernel $\kappa(x-x_n, y-y_n)$ to the sought after somewhat continuous value $I(x, y)$. Let S_k be the integrated signal provided by the kth detector (or dexel) centered on the location (ρ_k, θ_k) and $\eta(\rho-\rho_k, \theta-\theta_k)$ be the kernel modeling the measurement process. If no noise is assumed in the acquisition process, then the relation between S_k and I_n is

$$S_k = \sum_{n=1}^{N} I_n r_{k,n} \text{ with } r_{k,n} \text{ equal to} \tag{1}$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{0}^{\infty} \int_{-\pi}^{\pi} \eta(\rho - \rho_k, \theta - \theta_k)\kappa(x - x_n, y - y_n)\delta(x \cos \theta + y \sin \theta - \rho)\, \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}\rho\, \mathrm{d}\theta.$$

The discrete Radon matrix could be considered as a transition matrix whose coefficient $r_{k,n}$ are proportional to the probability of an activity in the nth pixel to be detected by the kth detector. All reconstruction algorithms make extensive use of this representation so that they can be viewed as algorithms dedicated to inversion of the relationship between the $\{I_n\}_{n \in \{1,...,N\}}$ and the $\{S_k\}_{k \in \{1,...,K\}}$ given by Eq. (1). Recovery of this probability prompts the need for post-normalization of expressions 1. The normalized projection operator is

$$S_k = \sum_{n=1}^{N} I_n p_{k,n} \quad \text{with} \quad p_{k,n} = \frac{r_{k,n}}{\sum_{n=1}^{N} r_{k,n}}.$$

The correct convergence of reconstruction algorithms towards the putative real discrete image is closely related to the suitability of the chosen kernels to provide an appropriate amount of smoothing. Oversmoothing tends to give smooth but blurred images while undersmoothing tends to amplify the effect of statistical fluctuations of observation and then to produce artefacts in the reconstructed images.

3. Framework and notations

Statistics applied to signal processing usually concern random observations, i.e. well defined processes leading to unpredictable outcomes [9]. This unpredictability is commonly assumed to be due to the randomness of the observed phenomenon and therefore associated with uncertainty. Uncertainty relates to the truth of an observation, understood as its conformity to reality. Imprecision is related to the content of the observation and therefore to the observation process itself [7]. For example, in emission tomography, uncertainty is related to random variation in the proportion of photons detected on a sensor during a finite time (in reference to the mean value of the Poisson process that rules the radioactive decay), and thus may vary from one experiment to another and depends on the object. The precision of the sensor is more related to its physical characteristics, such as energy or spatial resolution, and therefore does not depend on the observed object (for a given imaging device and a given photon energy).

3.1. Summative and non-summative kernels

One of the most common models of an uncertain observation is a probability density distribution, i.e. a summative kernel [17]. A summative kernel is a [0,1]-valued function κ defined on a domain Ω verifying the summative normalization property: $\int_{\Omega} \kappa(w) \, dw = 1$.

So-called kernel smoothing methods make use of this model to reduce the effect of quantization and noise by defining a weighted neighborhood of each sample [21]. This weighted neighborhood is then used in an additive aggregation process in order to provide an estimated value of the signal inside and outside the sampling grid.

In addition to this modeling, measurement errors can be considered as a feature of single observations due to systematic inaccuracy in the measurement process. This is a typical physicist's viewpoint where proper calibration of the sensor provides an estimate of the error domain (usually intervals) [14,26]. This domain is asked to certainly contain the measured information. This approach is used to evaluate the image of a function whose arguments are subsets. It is the most suitable modeling in cases where repeated observations cannot provide any additional information. This so-called bounded-error approach mainly concerns imprecision [14]. For both practical and theoretical reasons, validity domains of observations are usually intervals or boxes, i.e. Cartesian products of intervals.

The all-or-nothing nature of interval analysis, in contrast with the probability approach which accepts graduations, introduces asymmetry between them. The canonical generalization of interval analysis leads to the fuzzy-subset theory [7]. This fuzzy generalization of a validity domain relates to fuzzy measures [28].

Like a summative kernel, a non-summative kernel is a [0,1]-valued function π defined on a domain Ω verifying the maxitive normalization property:

$$Sup_{w \in \Omega} \pi(w) = 1.$$

π can be seen as a possibility distribution or as the membership function μ_F of a normalized fuzzy subset F of Ω.

3.2. Aggregation with Choquet integrals

Choquet integrals have become a central tool for aggregating values associated with non-additive confidence measures [16,10,5]. [1] In this section, we will introduce some of the properties of this aggregation operator that will be used in the sequel.

Let $\Phi = \{P_1, \ldots, P_N\}$ be a finite set of N elements. For all $n \in \{1, \ldots, N\}$, let I_n be a value associated with each element P_n. To begin, we suppose that all the values are positive. For the sequel we use the following notations:

Notation 1.
- $I : \Phi \to \mathbb{R}^+$ is the function defined by $I(P_n) = I_n$.
- We denote $(.)$ the permutation on $\{1, \ldots, N\}$ that sorts the values I_n in increasing order: $I_{(1)} \leqslant \cdots \leqslant I_{(N)}$.
- $A_{(n)}$ is the coalition of all elements whose value exceeds $I_{(n)}$: $A_{(n)} = \{P_i \in \Phi / I_i \geqslant I_{(n)}\} = \{P_{(n)}, \ldots, P_{(N)}\}$.

Definition 2.
- A capacity is a set function $v : \mathcal{P}(\Phi) \to [0, 1]$ such that $v(\emptyset) = 0$, $v(\Phi) = 1$, and
  $\forall A, B \in \mathcal{P}(\Phi) \; A \subseteq B \Rightarrow v(A) \leqslant v(B)$.
- $\bar{v}$ the conjugate capacity of v is defined by $\forall A \subseteq \Phi$, $\bar{v}(A) = 1 - v(A^c)$ where A^c is the complementary set of A.

Definition 3. A capacity v on Φ is
- a convex capacity if $\forall(A, B) \in \mathcal{P}(\Phi)$, $v(A) + v(B) \leqslant v(A \cup B) + v(A \cap B)$,
- a concave capacity if $\forall(A, B) \in \mathcal{P}(\Phi)$, $v(A) + v(B) \geqslant v(A \cup B) + v(A \cap B)$.

Note that if v is a concave capacity, then $\bar{v}$ is a convex capacity.

---

[1] Named capacities Vasilesco [25].

Definition 4. Let $\Phi = \{P_1, \ldots, P_N\}$ be a set of N elements, $I : \Phi \to \mathbb{R}^+$ be a function from Φ to the real positives and v be a capacity on $\mathcal{P}(\Phi)$.

- The Choquet integral of I with respect to the capacity v is:

$$C_v(I) = \sum_{n=1}^{N} I_{(n)}(v(A_{(n)}) - v(A_{(n+1)}))$$

$$= \sum_{n=1}^{N}(I_{(n)} - I_{(n-1)})v(A_{(n)}) \text{ with } A_{(N+1)} = \emptyset \text{ and } I_{(0)} = 0.$$

- A dual Choquet integral can be computed by using the conjugate capacity:

$$C_{\bar{v}}(I) = \sum_{n=1}^{N} I_{(n)}(v(A^c_{(n+1)}) - v(A^c_{(n)})).$$

Proposition 5. *If v is a concave capacity, then $C_v(I) \geqslant C_{\bar{v}}(I)$.*

Proof. We have $C_v(I) - C_{\bar{v}}(I) = \sum_{n=1}^{N}(I_{(n)} - I_{(n-1)})(v(A_{(n)}) - 1 + v(A^c_{(n)}))$. Since v is concave, $v(A_{(n)}) + v(A^c_{(n)}) \geqslant v(\Omega) + v(\emptyset) = 1$ which, when associated with $(I_{(n)} - I_{(n-1)}) \geqslant 0$, entails $C_v(I) - C_{\bar{v}}(I) \geqslant 0$. □

3.3. Aggregation with the asymmetric Choquet integral

When the range of I is $\mathbb{R}$, then, according to Denneberg [5], aggregation with respect to the capacity should be performed via the *asymmetric Choquet integral*.

In this subsection, we consider having a real function $I : \Phi \to \mathbb{R}$.

We denote as r the index belonging to $\{1, \ldots, N\}$, such that $I_{(1)} \leqslant \cdots \leqslant I_{(r)} < 0 \leqslant I_{(r+1)} \leqslant \cdots I_{(N)}$. Then we define $I^- = \{-I_{(1)}, \ldots, -I_{(r)}, 0, \ldots, 0\}$ and $I^+ = \{0, \ldots, 0, I_{(r+1)}, \ldots, I_{(N)}\}$.

Definition 6. Let $\Phi = \{P_1, \ldots, P_N\}$ be a set of N elements, $I : \Phi \to \mathbb{R}$ be a real function on Φ and v be a capacity on $\mathcal{P}(\Phi)$.

The asymmetric Choquet integral of I with respect to v is

$$\check{C}_v(I) = C_v(I^+) - C_{\bar{v}}(I^-).$$

Proposition 7. *If v is a concave capacity, then $\check{C}_v(I) \geqslant \check{C}_{\bar{v}}(I)$.*

This proposition is direct from Proposition 5.

4. Imprecise projection

In this section, we propose to show how it is possible to shift from additive summative kernel based to non-additive non-summative kernel based projection operators. In analogy to the summative approach, we go from a continuous binary to a weighted discrete projection.

4.1. Continuous case

Let us consider the image as a subset Ω of $\mathbb{R}^2$ where the continuous activity is considered to be binary. Each point of Ω is supposed to be active or not and its emission activity during the acquisition is taken as a unit. θ is the subset of all active points of Ω. The activity outside Ω is assumed to be unknown. The emission density of each subset A of Ω is assumed to be due to the local density of active points in A.

Each detector D is associated with a polar coordinate (ρ, θ) and an infinite projection zone W (see Fig. 1).

Fig. 1. Binary detected activity.

The activity S measured on detector D is assumed to be proportional to the surface of intersection of W with θ. The scale of the measured activity is assumed to be unknown. If the activity is a random process, then the measured activity S is also a random process. No additivity is hypothesized on the projection process.

Let S_1 be the projected activity on detector D_1 and S_2 be the projected activity on a disconnected detector D_2, the activity S_{12} on $D_1 \cup D_2$ is not assumed to be equal to $S_1 + S_2$. However, since the projection process is a measurement, the following relation has to be respected: $S_{12} \geqslant max(S_1, S_2)$.

Definition 8. The surface associated with an object on Ω is the function $|.|$ defined as follows:

$$|.| : \mathcal{P}(\Omega) \rightarrow \mathbb{R}$$

$$A \mapsto \int_A \mathrm{d}w,$$

where $\mathcal{P}(\Omega)$ denotes the class of subsets of Ω.

If S behaves like a density, then S is proportional to $|W \cap \theta|$. The basic projection formula in the continuous space should be $S = \zeta_W |W \cap \theta|$, where ζ_W is an unknown normalizing coefficient. Conventionally, we have: if $\theta = \Omega$ then $S = 1$ whatever the detector position, then $1 = \zeta_W |W \cap \Omega|$ and therefore: $S = |W \cap \theta|/|W \cap \Omega|$. Note that W is the (infinite) subset of $\mathbb{R}^2$ of any point of space that can project on detector D.

The measurement process consists of K projections S_k associated with K detector positions (or K different detectors) D_k, called the dexels. Each dexel D_k is associated with a polar coordinate (ρ_k, θ_k). S_k is assumed to be the proportion of photons emitted by the active zone that is collected by a detector associated with the polar coordinate during a time unit. The number of photons can be computed by multiplying S_k by the spatial interaction between the measurement and the image, i.e. $|W_k \cap \Omega|$ which provide the non-normalized projection, that is S is proportional to $|W \cap \theta|$.

4.2. Discretization of the image plane

With the representation of the image, always denoted Ω, being discrete, the continuous active zone is now imprecisely known. Fig. 2 shows a sampled version of the binary activity depicted in Fig. 1.

The non-summative kernel technique relies on a set-theoretic based interplay between continuous and discrete involving the use of tiling and partitioning. In fact, there are three kinds of binary regular discretization that can be viewed in Fig. 3.

Definition 9. Let $\{C_n\}_{n \in \{1,...,N\}}$ be a covering of a plane image Ω.

• $\{C_n\}_{n \in \{1,...,N\}}$ is a partition of Ω if and only if

$$\bigcup_{n=1}^{N} C_n = \Omega, \tag{2}$$

$$C_n \cap C_m = \emptyset \quad \text{if } n \neq m. \tag{3}$$

Fig. 2. Discrete active zone.

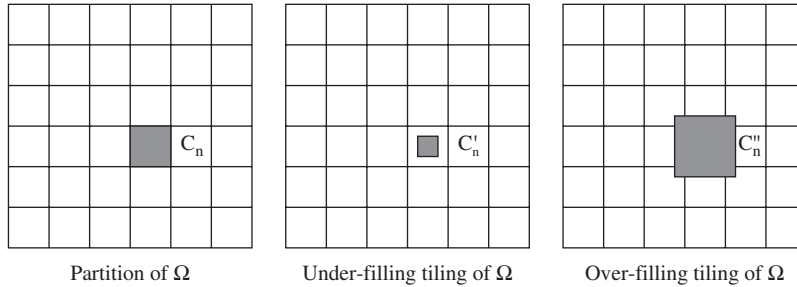| Partition of Ω | Under-filling tiling of Ω | Over-filling tiling of Ω |

Fig. 3. Discretization of the plane.

- $\{C_n\}_{n \in \{1,\dots,N\}}$ is an under-filling tiling if it follows property (3) but not property (2).
- $\{C_n\}_{n \in \{1,\dots,N\}}$ is a over-filling tiling if it follows property (2) but not property (3).

Note that the first property seems to be necessary to ensure a continuous to discrete transition without loss and the second one ensures additivity.

Because of discretization, the only available information about the active zone is a measure of the discrete density I_n via an unknown kernel associated with pixel P_n. The notion of continuous zone of binary activity has to be replaced by a discrete coalition at a given level.

As done in the continuous case, we consider K dexels D_k. For each dexel, we want to model S_k, the detected activity associated with dexel D_k.

Let $\Phi = \{P_1, \dots, P_N\}$ [2] be the set of N pixels of the image. Each discrete entity P_n is associated with a subset C_n of Ω.

First, we suppose that all I_n to be non-negative. The intensity is not a binary value so we need to define an intensity level.

Definition 10. Let γ be a real value, θ_γ is the subset of pixels whose intensity exceeds γ:

$$\theta_\gamma = \{P_n \in \Phi | I_n \geqslant \gamma\}.$$

Note that $\theta_0 = \Phi$.

Now by analogy with the continuous case, we define W_k as the discrete subset of pixels that can potentially cause the detected density S_k on the kth dexel D_k. This subset of $\mathbb{R}^2$ defines a binary possibility distribution associated with each D_k denoted $\pi(D_k; .)$, and defined as follows:

$$\forall w \in \Omega, \quad \begin{cases} \pi(D_k; w) = 1 & \text{if } w \in W_k, \\ \pi(D_k; w) = 0 & \text{else.} \end{cases}$$

---

[2] To be clear, we already named Φ as the finite set of elements to define the Choquet integral in Section 3.2.

This continuous possibility distribution on $\mathbb{R}^2$ induces a discrete possibility distribution on Φ:

$$\pi(D_k; P_n) = \bigvee_{w \in C_n} \pi(D_k; w).$$

Consequently, $\Pi(D_k; P_n) = 1$ if there is $w \in C_n$ such that $\pi(D_k; w) = 1$. In the following $\Pi(D_k; P_n)$ is denoted $\pi_{k,n}$.

Dually, $\pi_{k,n}$ defined on Φ a subset $W_k = \{P_n \in \Phi | \Pi(D_k, P_n) = 1\}$.

To define the capacity used in the following we have to express the function $|.|$ which is the surface according to the continuous case.

Definition 11. Let Ω be a plane image, we denote Φ the finite set of pixels and $\{C_n\}_{n \in \{1,\ldots,N\}}$ a covering, then

$$|.| : \Phi \to \mathbb{R},$$
$$A \mapsto |\cup_{n / P_n \in A} C_n|.$$

Proposition 12. S_k, the detected density associated with the kth dexel D_k, is the Choquet integral of density $I : \Phi \to \mathbb{R}^+$ with respect to the capacity v_k defined by

$$\forall A \subseteq \Phi, \quad v_k(A) = \frac{|W_k \cap A|}{|W_k \cap \Phi|}. \tag{4}$$

Proof. For each threshold $\gamma \in \mathbb{R}$, a binary detected intensity can be computed using $S_k^\gamma = \zeta_k |W_k \cap \theta_\gamma|$. ζ_k is an unknown normalization factor, and $|W_k \cap \theta_\gamma|$ is the surface on the image associated with the subset of pixels $W_k \cap \theta_\gamma$. To be more precise, we have $|W_k \cap \theta_\gamma| = |\bigcup_{n | P_n \in W_k \cap \theta_\gamma} C_n|$.

To obtain S_k, the detected intensity associated with the kth dexel D_k, we aggregate all the S_k^γ:

$$S_k = \zeta_k \int_0^\infty S_k^\gamma \, d\gamma = \zeta_k \int_0^\infty |W_k \cap \theta_\gamma| \, d\gamma.$$

The normalization factor can be obtained by assuming the usual normalization: if $\forall n$, $I_n = 1$, then $\forall k$, $S_k = 1$. So we have $S_k = \zeta_k \int_0^\infty S_k^\gamma \, d\gamma = \zeta_k |W_k \cap \theta_0| 1 = 1$ and therefore $\zeta_k = 1/|W_k \cap \theta_0|$ which entails $S_k = \int_0^\infty (|W_k \cap \theta_\gamma|/|W_k \cap \theta_0|) \, d\gamma$. Since the possible values of γ are discrete, the above formula can be computed by

$$S_k = \sum_{n=1}^N (I_{(n)} - I_{(n-1)}) \frac{|W_k \cap \theta_{I_{(n)}}|}{|W_k \cap \theta_0|},$$

where $(.)$ is the permutation which sorts the pixels according to increasing intensity, which concludes the proof. $\quad \square$

Proposition 13. Let $\{C_n\}_{n \in \{1,\ldots,N\}}$ be a covering of a plane image Ω,

- if $\{C_n\}_{n \in \{1,\ldots,N\}}$ is a partition or an under-filling tiling then v_k is additive and $S_k = \sum_{n=1}^N I_n v_k(P_n)$,
- if $\{C_n\}_{n \in \{1,\ldots,N\}}$ is an over-filling tiling, then v_k is concave.

Proof.

- It is trivial to show the first assertion because, if $\{C_n\}_{n \in \{1,\ldots,N\}}$ is a partition or an under-filling tiling, then $C_n \cap C_m = \emptyset$ if $n \neq m$, which entails the additivity of the capacity v_k.
- For the second assertion, we have to prove that

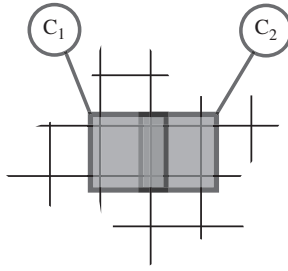$$\forall (A, B) \subseteq \Phi^2, \ v_k(A) + v_k(B) \geqslant v_k(A \cup B) + v_k(A \cap B).$$

Fig. 4. Over-filling tiling.

First recall that the function $|.|$ defines the surface associated with an object on Ω. So as we have $v_k(A) = (|W_k \cap A|/|W_k \cap \Phi|) = \zeta_k |W_k \cap A|$.

We need to compute $|W_k \cap A|$ with the over filling tiling $\{C_n\}_{n \in \{1,\ldots,N\}}$. To do this we have to define the following set:

$$\mathcal{A} = \bigcup_{\{n \mid P_n \in W_k \cap A\}} C_n, \quad \mathcal{B} = \bigcup_{\{n \mid P_n \in W_k \cap B\}} C_n,$$

$$\mathcal{AB} = \bigcup_{\{n \mid P_n \in W_k \cap (A \cap B)\}} C_n, \quad \mathcal{A} \cup \mathcal{B} = \bigcup_{\{n \mid P_n \in W_k \cap (A \cup B)\}} C_n.$$

So we have $v_k(A) = \zeta_k |\mathcal{A}|$, $v_k(B) = \zeta_k |\mathcal{B}|$, $v_k(A \cap B) = \zeta_k |\mathcal{A} \cap \mathcal{B}|$, and $v_k(A \cup B) = \zeta_k |\mathcal{A} \cup \mathcal{B}|$. As $\{C_n\}$ is over-filling tiling, $\mathcal{AB} \subseteq \mathcal{A} \cap \mathcal{B}$ and thus $|\mathcal{AB}| \leqslant |\mathcal{A} \cap \mathcal{B}|$. Since $|\mathcal{A}| + |\mathcal{B}| = |\mathcal{A} \cap \mathcal{B}| + |\mathcal{A} \cup \mathcal{B}|$, $v_k(A) + v_k(B) = \zeta_k |\mathcal{A} \cap \mathcal{B}| + \zeta_k |\mathcal{A} \cup \mathcal{B}|$, and therefore $v_k(A) + v_k(B) \geqslant v_k(A \cup B) + v_k(A \cap B)$.

Note that, as the previous inequality can be strict, the capacity v_k is not additive. To illustrate a strict inequality, let us consider two contiguous pixels $P1$ and $P2$ associated with an over-filling tiling $\{C1, C2\}$ (see Fig. 4).

We have $v_k(P1) = |C1| = |C2| = v_k(P2)$, $v_k(P1 \cap P2) = v_k(\emptyset) = 0$ and $v_k(P1 \cup P2) = |C1 \cup C2| < |C1| + |C2|$ which implies $v_k(P1 \cup P2) < v_k(P1) + v_k(P2)$. □

As shown in Section 3.2, with v_k being a concave capacity, an upper estimation $\overline{S_k}$ and a lower estimation $\underline{S_k}$ can be defined for each dexel D_k.

Definition 14. Let Φ be the set of pixels, and I be the positive discrete density. For the kth dexel D_k,

- The upper projection is

$$\overline{S_k} = \sum_{n=1}^{N} I_{(n)}(v_k(A_{(n)}) - v_k(A_{(n+1)})) = \sum_{n=1}^{N} (I_{(n)} - I_{(n-1)})v_k(A_{(n)}). \tag{5}$$

- The lower projection is

$$\underline{S_k} = \sum_{n=1}^{N} I_{(n)}(\bar{v}_k(A_{(n)}) - \bar{v}_k(A_{(n+1)})) = \sum_{n=1}^{N} (I_{(n)} - I_{(n-1)})\bar{v}_k(A_{(n)}) \tag{6}$$

with $v_k(A) = (|W_k \cap A|/|W_k \cap \Phi|) \ \forall A \subseteq \Phi$ where $W_k = \{P_n \in \Phi | \Pi(D_k, P_n) = 1\}$.

4.3. Weighted possibility distribution

If the possibility distribution is weighted, in order to provide a good trade-off between enhancing and blurring details [23], the formula of the lower and upper projection has to be modified. As the difference between the two expressions is the use of the conjugate capacity, we only give the proof for the upper estimation.

For any confidence level α, the quantity $\overline{S_k^{\alpha}}$ can be computed by using Eq. (5) with the capacity defined by $v_k^{\alpha}(A_{(n)}) = \zeta_k |W_k^{\alpha} \cap \theta_{I_{(n)}}|$. So we have

$$\overline{S_k^{\alpha}} = \zeta_k \sum_{n=1}^{N}(I_{(n)} - I_{(n-1)})|W_k^{\alpha} \cap \theta_{I_{(n)}}|,$$

where ζ_k is an unknown normalization factor, W_k^{α} is the set of pixels whose interaction with the kth dexel exceeds the level α: $W_k^{\alpha} = \{P_n | \pi_{k,n} \geqslant \alpha\}$.

The aggregation of all projected values $\overline{S_k^{\alpha}}$ can be obtained by integrating according to the confidence levels with an a-posteriori normalization:

$$\overline{S_k} = \zeta_k \int_0^1 \overline{S_k^{\alpha}}\, d\alpha$$

$$= \zeta_k \int_0^1 \left(\sum_{n=1}^{N}(I_{(n)} - I_{(n-1)})|W_k^{\alpha} \cap \theta_{I_{(n)}}| \right) d\alpha$$

$$= \zeta_k \sum_{n=1}^{N} \left((I_{(n)} - I_{(n-1)}) \int_0^1 |W_k^{\alpha} \cap \theta_{I_{(n)}}|\, d\alpha \right).$$

Normalization is obtained as usual by assuming $\overline{S_k} = I$ if: $I_n = I\ \forall n$; therefore

$$\overline{S_k} = \sum_{n=1}^{N}(I_{(n)} - I_{(n-1)})v_k(A_{(n)}) \quad \text{with } v_k(A) = \dfrac{\displaystyle\int_0^1 |W_k^{\alpha} \cap A|\, d\alpha}{\displaystyle\int_0^1 |W_k^{\alpha} \cap \Phi|\, d\alpha}.$$

The corresponding lower projection can be obtained by using the conjugate capacity of v_k.

4.4. Properties of upper and lower projections

In most reconstruction algorithms, projection operators are iteratively used to project and back-project the reconstruction error until convergence. Thus, the value to be projected is generally not a non-negative value and therefore the usual Choquet integral has to be replaced by an asymmetric Choquet integral, as presented in Section 3.3, for upper and lower projection computations.

Definition 15. Let Φ be the set of pixels, and I be the discrete density. For the kth dexel D_k,

- The upper projection is: $\overline{S_k} = \check{C}_{v_k}(I)$,
- The lower projection is: $\underline{S_k} = \check{C}_{\bar{v}_k}(I)$.

The upper and lower projection properties are directly inherited from the properties of Choquet integrals and the most interesting property is derived from the work of Schmeidler [19] and concerns the core of the convex capacities.

Definition 16. For all convex capacities we can define the core of the capacity as follows:

$$core(v) = \{p \text{ probalility such that } v(A) \leqslant p(A),\ \forall A\}.$$

Note that, if p belongs to $core(v)$, then $\forall A \in \Omega$, $v(A) \leqslant p(A) \leqslant \bar{v}(A)$.
We have the following results due to Schmeidler and Dellacherie [19,8,4].

Proposition 17. *Let Φ be a finite set, $I : \Phi \to \mathbb{R}$ be a function and v be a convex capacity on Φ. Then,*

$$\check{C}_v(I) = Inf_{p \in core(v)} \int I \, dp,$$

$$\check{C}_{\bar{v}}(I) = Sup_{p \in core(v)} \int I \, dp.$$

When I is a discret function which the values are denoted $I_1, \ldots, I_N$, for any probability p, $\int I \, dp = \sum_{i=1}^{N} I(n) p(P_n)$.

Our capacity v_k defined by Eq. (4) is concave, so the above result leads us to the final property.

Proposition 18.

$$\overline{S_k} = Sup_{p \in core(\bar{v}_k)} \sum_{n=1}^{N} p(P_n) I_n$$

$$\underline{S_k} = Inf_{p \in core(\bar{v}_k)} \sum_{n=1}^{N} p(P_n) I_n.$$

Since usual Radon matrices are defined by summative kernels, $\overline{S_k}$ (rsp. $\underline{S_k}$) is an upper (rsp. lower) bound of any conventional projection obtained with a Radon matrix whose coefficients can be derived from a probability function which belongs to the core of $\bar{v}_k$. Therefore, the distance between the non-additive aggregated values obtained by using v_k and $\bar{v}_k$ reflects the lack of stability of a conventional aggregation process when there is variation in the hypothesized confidence distribution. Since this stability depends on the robustness of the hypothesized model and the local noise level, it can be used as a measure of aggregation error.

5. Numerical implementation

The only difficulty in computing upper and lower projections concerns the computation of the capacity v_k defined by (4). The general computation of the Choquet integral involves an iterative estimation of capacities of nested sets. The easiest way to perform this iterative estimation is first to invert the sum in expression: $\overline{S_k} = \sum_{n=1}^{N} (I_{(n)} - I_{(n-1)}) v_k(A_{(n)})$. As we have $A_{(n)} = A_{(n+1)} \cup P_{(n)}$, the value $v_k(A_{(n)})$ can be deduced from the value of $v_k(A_{(n+1)})$ by using a map-based algorithm. This decomposition requires a non-summative sampling kernel. As shown in [6], any summative kernel with support $[-\delta, \delta] \times [-\delta, \delta]$ is dominated by the binary kernel with support $[-\delta, \delta] \times [-\delta, \delta]$; and more specifically by the fuzzy kernel with support $[-\delta, \delta] \times [-\delta, \delta]$ with a pyramidal membership function.

We first present this algorithm when the partition $\{C_n\}_{n \in \{1,\ldots,N\}}$ is binary. Next we present it when the partition $\{C_n\}_{n \in \{1,\ldots,N\}}$ is fuzzy.

(1) *Binary kernel*: the general shape of C_n is depicted in Fig. 5.
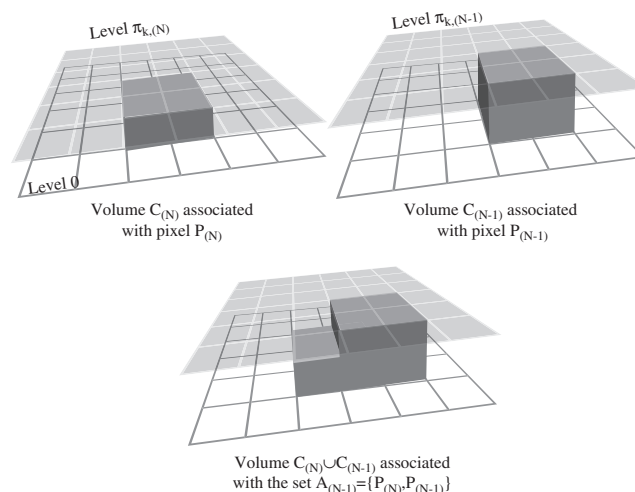

Fig. 5. Iterative computation of the capacity.

Let N_l be the number of lines and N_r be the number of rows in the image. The algorithm needs a matrix $T(l, r)$ whose size is $(N_l + 1) \times (N_r + 1)$. Let $(l_{\{n\}}, r_{\{n\}})$ be the image coordinate of pixel P_n. The iterative computation of the $v_k(A_{(n)})$ works with algorithm 1:

Data: matrix T
Result: the capacity v_k
Begin

Initialization of T as a null matrix. **for** $n = N$ *down to 1* **do**

for $u = l_{(n)}$ *to* $l_{(n)} + 1$ **do**

for $v = R_{(n)}$ *to* $R_{(n)} + 1$ **do**

t(u,v)= MAX(t(u,v), $\Pi_{k,n}$)

$$v_k(A_{(n)}) = \delta^2 \sum_{u,v} T(u, v)$$

End

Algorithm 1: Binary kernel.

(2) *Fuzzy kernel*: Matrix T requires a third coordinate: $T(l, r, s)$, where $s \in \{1, 2, 3, 4\}$. Each coordinate corresponds to a truncated quarter of pyramid, as depicted in Fig. 6.
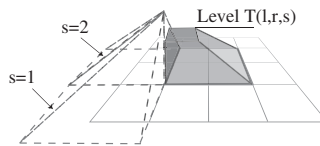

Fig. 6. Elementary truncated quarter of a pyramid.

The iterative computation works with algorithm 2:

Data: matrix T
Result: the capacity v_k
begin

Initialization of T as a null matrix. **for** $n = N$ *down to 1* **do**

s=1 **for** $u = l_{(n)}$ *to* $l_{(n)} + 1$ **do**

for $v = R_{(n)}$ *to* $R_{(n)} + 1$ **do**

t(u,v,s)= MAX(t(u,v,s), $\Pi_{k,n}$) s = s + 1

$$v_k(A_{(n)}) = \delta^2 \sum_{u,v,s} T(u, v, s)$$

end

Algorithm 2: Fuzzy kernel.

Computation of $v_k(A_{(n)})$ involves the geometrical computation of a union of four truncated quarters of pyramids that can be easily achieved but whose algorithm is too long to present in this paper.

6. Experimentations: ability to estimate projected noise

In this section, we illustrate the ability of the upper and lower projection operators to estimate the local noise.

For the experiment, SPECT projections of an uniform cylinder and from a Jaszczak resolution phantom (JRP) were acquired after filling these phantoms with 740 MBq (20 mCi) of a technetium 99 m solution. On average, 507 and 381 kcounts were acquired for each line of the sinogram used for the cylinder and the JRP, respectively. One hundred
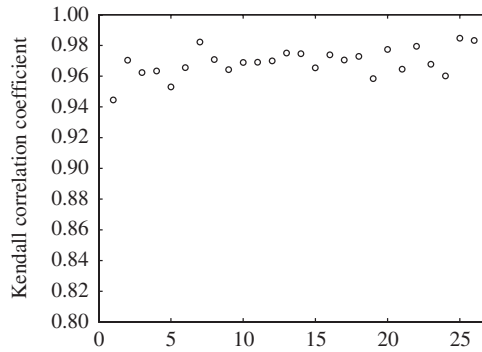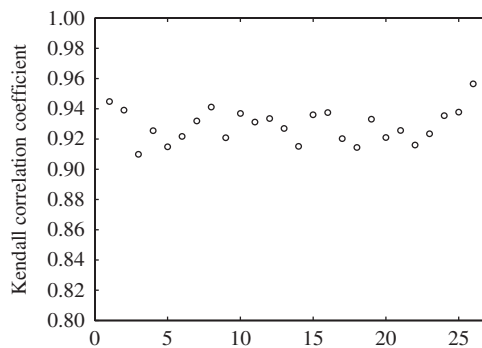
Fig. 7. Resolution phantom.

Fig. 8. Uniform cylindrical distribution.

and twenty-eight projections over a 360° arc were recorded. Twenty-five acquisitions of the cylindrical phantom were used to record projections from a uniform cylindrical radioactivity distribution and 25 acquisitions of the JRP were used as a resolution phantom. Each of the 25 acquisitions for the two phantoms is supposed to be a repetition of the same event. For each one of the phantoms, 25 slices were computed using the usual back-projection procedure with the probability coefficient associated with a uniform kernel. Concurrently, lower and upper backprojections were computed as described above.

For a given backprojected pixel P_n, a variance v_n was computed using it within each of the 25 backprojections. Then, for the same pixel P_n, a length l_n^s of the confidence interval was computed for each of the backprojected slices $s \in \{1, \ldots, 25\}$. To be more precise, for each s and each n, l_n^s is the difference between the upper and the lower back-projections.

We will now test whether the distribution of the estimated variances $\{v_n\}_{n \in \{1,\ldots,N\}}$ is correlated or not with the 25 distributions of the length $\{l_n^s\}_{n \in \{1,\ldots,N\}}$, i.e. if this length can be regarded as a spread factor measuring the local noise level. In order to perform a distribution free-comparison, we use the Kendall correlation coefficient, since it is known to be one of the most powerful and intuitively simple measurement of the strength of relationship between two variables [20].

We compute, for two experiments and for each of the 25 acquisitions, the values of the Kendall correlation coefficient between the distribution of the length versus the estimated variance for each backprojected slice. These values are shown in Figs. 7 and 8.

For both experiments, the correlation coefficient values are such that the independence between predicted spread and variance has to be rejected (all the p-values are lower than 10^{-8} for the first experiment and lower than 10^{-3} for the second experiment).

7. Conclusion

Here, we have proposed a new operator to perform discrete Radon-like projections (and back-projections) in emission tomography. These operators are based on non-additive aggregation methods using non-summative kernel-based Radon matrices. A consequence of using such a method is that the projected (or backprojected) values are imprecise, i.e. known as an interval. The length of this interval varies according to the spread of all values obtained using conventional approaches with different conventional Radon matrices. We have carried out theoretical and experimental studies to highlight the properties of the interval provided by this new operator and particularly the correlation of its length with the random variation of the image under reconstruction.

In the experiment, we have shown that the difference between the upper and lower back-projection of a sinogram is highly correlated with the variance of the distribution of its conventional back-projections.

In a future study, we would like to utilize the potential of this new operator to project the error and propose new reconstruction methods to obtain a reconstructed image whose confidence interval is known. Other studies must be carried out to design proper statistical tools to handle these intervals. Such methods are highly connected to imprecise probability approach [27]. This will be done prior to clinical validation studies.

Acknowledgment

References

[1] A. Aldroubi, M. Unser, M. Eden, Fast b-spline algorithms for continuous image representation and interpolation, IEEE Trans. Pattern Anal. Mach. Intell. 13 (1991) 277–285.

[2] H. Barrett, D. Wilson, B. Tsui, Noise properties of the em algorithm: I. Theory, Phys. Med. Boil. 39 (1994) 833–846.

[3] G. Choquet, Theory of capacities, Ann. l'Inst. Fourier 5 (1953) 131–295.

[4] C. Dellacherie, Quelques commentaires sur les prologements de capacités, in: Séminaire de probabilités, vol. 5, Strasbourg, 1971, pp. 77–81.

[5] D. Denneberg, Non-Additive Measure and Integral, Kluwer, Dordrecht, 1994.

[6] D. Dubois, L. Foulloy, G. Mauris, H. Prade, Probability-possibility transformations, triangular fuzzy sets, and probabilistic inequalities, Reliab. Comput. 10 (2004) 273–297.

[7] D. Dubois, H. Prade, Possibility Theory: An Approach to Computerized Processing of Uncertainty, Plenum Press, London, 1988.

[8] T. Fetz, M. Oberguggenberger, Propagation of uncertainty through multivariate functions in the framework of sets of probability measures, Reliab. Eng. Syst. Saf. 85 (1–3) (2004) 73–87.

[9] G. Gillen, A simple method for the measurement of local statistical noise level in spect, Phys. Med. Biol. 37 (1992) 1573–1579.

[10] M. Grabisch, T. Murofushi, M. Sugeno, Fuzzy Measures and Integrals: Theory and Applications, Physica-Verlag, Berlin, 2000.

[11] J. Guédon, Y. Bizais, Spline-Based Regularization for Discrete FBP Reconstruction, in: Lecture Notes in Computer Science, IPMI Springer, 1991, pp. 55–64.

[12] J. Guédon, Y. Bizais, Bandlimited and haar filtered back-projection reconstructions, IEEE Trans. Med. Imaging 13 (1994) 430–440.

[13] F. Jacquey, K. Loquin, F. Comby, O. Strauss, Non-additive approach for gradient-based detection, in: ICIP'07, San Antonio, TX, 2007, pp. 49–52.

[14] L. Jaulin, M. Kieffer, O. Didrit, E. Walter, Applied Interval Analysis, Springer, New York, 2001.

[15] A. Kak, M. Slaney, Principles of computerized tomographic imaging, in: Society for Industrial and Applied Mathematics, Roberto'Malley, USA, 2001.

[16] J.L. Marichal, An axiomatic approach of the discrete choquet integral as a tool to aggregate interacting criteria, IEEE Trans. Fuzzy Syst. 8 (2000) 800–807.

[17] E. Parzen, Non parametric statistical data modeling, J. Amer. Statist. Assoc. 74 (1979) 105–131.

[18] J. Qi, Resolution and noise properties of map reconstruction for fully 3d pet, IEEE Trans. Med. Imaging 19 (2000) 493–506.

[19] D. Schmeidler, Integral representations without additivity, Proc. American Mathematical Society 97 (1986) 255–261.

[20] S. Siegel, Non Parametric Statistics, McGraw Hill, New York, 1988.

[21] B. Silverman, Density Estimation for Statistics and Data Analysis, Chapman & Hall/CRC, London, 1986.

[22] J. Stayman, J. Fessler, Efficient calculation of resolution and covariance for penalized-likelihood reconstruction in full 3d spect, IEEE Trans. Med. Imaging 23 (2004) 1543–1556.

[23] O. Strauss, D. Mariano-Goulart, Reduction of sampling artefacts in single photon emission computed tomography using fuzzy-cell hough transform, Technical Report, LIRMM, 2006.

[24] M. Unser, A. Aldroubi, M. Eden, B-spline signal processing: Part I—theory, IEEE Trans. Signal Process. 41 (1993) 821–833.

[25] F. Vasilesco, La notion de capacité, Hermann, Paris, 1937.

[26] R. Viertl, Statistical Methods for Non-Precise Data, CRC Press, FL, 1996.

[27] P. Walley, Statistical Reasoning with Imprecise Probabilities, Chapman & Hall, New York, 1991.

[28] R. Yager, Uncertainty representation using fuzzy measures, IEEE Trans. Syst. B 32 (1) (2002) 13–20.