



OPEN

Quality measures for fully automatic CT histogram-based fat estimation on a corpse sample

Sebastian Schenkl^{1,3}, Michael Hubig^{1,3}✉, Holger Muggenthaler¹, Jayant Subramaniam Shanmugam¹, Felix Güttler², Andreas Heinrich², Ulf Teichgräber² & Gita Mall¹

In a previous article a new algorithm for fully automatic 'CT histogram based Fat Estimation and quasi-Segmentation' (CFES) was validated on synthetic data, on a special CT phantom, and tested on one corpse. Usage of said data in FE-modelling for temperature-based death time estimation is the investigation's number one long-term goal. The article presents CFES's results on a human corpse sample of size $R = 32$, evaluating three different performance measures: the τ -value, measuring the ability to differentiate fat from muscle, the anatomical fat-muscle misclassification rate D , and the weighted distance S between the empirical and the theoretical grey-scale value histogram. CFES-performance on the sample was: $D = 3.6\%$ for weight exponent $\alpha = 1$, slightly higher for $\alpha \geq 2$ and much higher for $\alpha \leq 0$. Investigating τ , S and D on the sample revealed some unexpected results: While large values of τ imply small D -values, rising S implies falling D and there is a positive linear relationship between τ and S . The latter two findings seem to be counter-intuitive. Our Monte Carlo analysis detected a general umbrella type relation between τ and S , which seems to stem from a pivotal problem in fitting Normal mixture distributions.

Abbreviations

| | |
|---------------------|---|
| M_j | Material with index j ($1 \leq j \leq J$) |
| Ξ | Index for materials of interest: $\Xi = F$ for 'fat' and $\Xi = M$ for 'muscle' |
| G | $\{0, \dots, N-1\}$ set of the N possible grey-scale values in CT slice |
| G_0 | Interval in G containing $E_0 = 0$ HU and approximately all grey-scale values n in G indicating materials F or M |
| Q | Number of pixels in CT slice |
| Q_Ξ | Number of pixels of material Ξ in CT slice |
| \underline{Q} | $\{1, \dots, Q\} := \{1, \dots, Q_F, Q_F + 1, \dots, Q_M + Q_F\}$ index set of all $Q = Q_F + Q_M$ pixels in a slice. The values Q_Ξ vary from slice to slice |
| \underline{Q}_Ξ | $\{1, \dots, Q_\Xi\}$ index set of Ξ -pixels in slice with number Q_Ξ of Ξ -pixels after suitable renumbering with $\Xi = F, M$. The sets \underline{Q}_Ξ vary from slice to slice |
| Y | (y_1, \dots, y_Q) Single CT slice with grey-scale values y_q at the $q = 1, \dots, Q$ pixels in G |
| Y_Ξ | $(y_{\Xi,1}, \dots, y_{\Xi,Q})$ partial CT slice of Ξ -pixels in slice Y |
| Ψ | $(\underline{Y}_1, \dots, \underline{Y}_R)$ sample of CT scans \underline{Y}_r of R bodies' abdomina with $r = 1, \dots, R$ |
| \underline{Y}_r | $(Y_{r,1}, \dots, Y_{r,L(r)})$ sequence of slices $Y_{r,l}$ with $l = 1, \dots, L(r)$ of abdominal scan \underline{Y}_r |
| $Y_{r,l}$ | $(Y_{r,l,1}, \dots, Y_{r,l,Q})$ sequence of pixel grey-scale values $y_{r,l,q}$ in G of slice $Y_{r,l}$ with $q = 1, \dots, Q$ |
| $Y_{\Lambda,r}$ | Series ($r = 1, \dots, R$) of simulated CT slices: $z = 0.5 < = > : \Lambda = A$; $z = 0.3 < = > : \Lambda = B$ |
| E_Ξ | Expectation value of Ξ -pixel grey-scale value in one slice $Y_{r,l}$ |
| S_Ξ | Standard deviation of Ξ -pixel grey-scale value in one slice $Y_{r,l}$ |
| z | Q_F/Q ratio of fat pixels or fat ratio in one slice $Y_{r,l}$ |
| RE_z | $ z^* - z /z^*$ relative error of estimator z^* |
| Θ | (E_F, E_M, S_F, S_M, z) parameter vector of interest |
| Φ | $\Phi(y, E, S)$ Normal distribution pdf evaluated at grey-scale value y in G with expectation value E and standard deviation S |
| f_Ξ | $f_\Xi(y, \Theta) = \Phi(y, E_\Xi, S_\Xi)$ grey-scale value pdf on G of material Ξ evaluated at grey-scale value y |

¹Institute of Legal Medicine, Jena University Hospital, 07749 Jena, Germany. ²Institute for Diagnostic and Interventional Radiology, Jena University Hospital, 07749 Jena, Germany. ³These authors contributed equally: Sebastian Schenkl and Michael Hubig. ✉email: michael.hubig@med.uni-jena.de

| | |
|--------------|---|
| f | $f(y, \Theta) = z f_F(y, \Theta) + (1-z) f_M(y, \Theta)$ overall grey-scale value pdf on G evaluated at grey-scale value y |
| X | $X(Y) = (x_0(Y), \dots, x_{N-1}(Y))$ empirical grey-scale value histogram of slice Y |
| X_{Ξ} | Partial histogram of material Ξ in slice with: $X = X_F + X_M$ |
| $X(\Theta)$ | Histogram computed by $X(\Theta) := Q f(G, \Theta)$ |
| τ | $\tau(Y)$ fat-muscle separability |
| S | $S_{\alpha}(X(Y), X(\Theta))$ WLS distance between $X(Y)$ and $X(\Theta)$ with weight exponent α |
| I_j | Interval in G with: For all pixels indices q : Estimation of quasi-segmentation: Pixel q associated with material $M_j \Leftrightarrow y_q$ in I_j |
| a_j | Constant grey value associated with material M_j for quasi-segmentation |
| $W(Y)$ | Quasi-segmentation result for slice Y : Pixel image of same format as Y with: $w_q := a_j \Leftrightarrow y_q$ in $I_j \Leftrightarrow$ pixel in $W(Y)$ with index q associated with material M_j |
| d | $d_{\alpha}(Y, W^*(Y))$ FM misclassification of slice Y ($d = 0$ correctly classified, $d = 1$ misclassified) |
| D | $D_{\alpha}(\underline{Y}, W^*(\underline{Y}))$ mean FM misclassification over all slices Y of scan \underline{Y} |
| F | $F_{\alpha}(\Psi, W^*(\Psi))$ mean sample FM misclassification over all scans \underline{Y} of sample Ψ |
| t_{FM} | Threshold fat/muscle in G |
| v | CT slices' voxel volume |
| ρ_{Ξ} | Mass density of material Ξ |
| m_{Ξ} | $V \rho_{\Xi} Q_{\Xi}$ mass of material Ξ |

In forensic temperature based death time estimation (TDE) casework rectal temperature measurement is the method of choice, the geometry and material parameters of the measurement locus's actual surrounding are important for modelling the cooling process. Its low conductivity (see e.g.¹ Table 6.23) makes body fat quantity and localization a potentially important issue in modelling post mortem temperature decline. The Finite Element Method (FEM) for TDE was developed to compute physics-based body core cooling. Using a CT scan to generate an FE mesh^{2,3} seems promising. One step is segmentation: matching a 'material label' to each of the CT scans pixels. The CT grey-scale values of body fat from muscle- and organ tissue are similar (see e.g.^{2,3}), making segregating challenging. In⁴ we described an algorithm "CT histogram-based Fat Estimation and quasi-Segmentation" (CFES) for fat—muscle segmentation which was applied to only one human corpse and validated against synthetic and CT phantom data. The actual paper now presents CFES results on a sample of $R = 32$ CT scans of human corpses.

All three quality measures used, the τ -value (signal theory based), the anatomical misclassification rate D , and the distance measure S applied to the grey-scale value histogram, depend on CFES's weighting exponent α (see⁴) directly as S or indirectly, as τ and D do.

Methods

The study was reviewed and approved by the ethics committee of the University Hospital Jena. All of our experiments were performed in accordance with the relevant guidelines and regulations (see subsection 'Ethics approval' below).

Fat tissue (**fat**, abbreviated 'F') has an expectation value of the CT attenuation equivalent below the water value (0 HU), while the value of muscle-, organ-, and connective tissues (**muscle** or 'M') is slightly above 0 HU (see e.g.⁵).

We will use the assumption that to any pixel in every slice a material type M_i can be assigned unambiguously accepting a small amount of misclassification. The term **FM frequency** denotes the number of pixels of a fixed grey-scale value in a CT slice.

In statements, true for fat as well as for muscle, we write Ξ as a symbol to be substituted by the letter 'F' for fat or by an 'M' for muscle. For any finite set D the symbol $\#D$ means the number of elements in D . For any function f and any set B in the preimage of f , the term $f(B)$ stands for the set of images $f(b)$ of elements b in B . The asterisk '*' attached to a quantity-symbol means the estimated value of the quantity, whereas an attached cross '+' means the true value of the indexed quantity. Generally symbols without '*' or '+' index are random variables. If dependence of a random variable on the real number parameter α is emphasized, the variable is indexed by ' α '. For any quantity U —which is interpreted as a random variable—and any set V in the image of U the symbol $\mu_V(U)$ stands for the mean of the quantity U , where the variation of U is bounded to the set V of possible U -values. For our list of symbols we assume for conciseness, that the slice contains fat and muscle pixel only. This is not the case in reality which is considered by constraining to the image parts where F and M are the only components. We use the abbreviations LS for Least Squares and WLS for Weighted Least Squares to specify estimation procedures.

For conciseness we will omit arguments or indices of quantity symbols if the omitted parts are unambiguously clear from the context.

We suppose the random variables y_q for $q = 1, \dots, Q$ in a slice Y to be stochastically independent.

The problem of fat-muscle quasi-segmentation, called **FM quasi-segmentation** in the following, makes it necessary to distinguish the partial histograms X_F and X_M of the grey-scale value histogram $X(Y)$ of a slice Y . As X_F and X_M are connected to the grey-scale value distributions f_F and f_M which are amalgamated to $f = z \cdot f_F + (1-z) \cdot f_M$ we chose the approach to firstly estimate z , f_F and f_M from the data. This enables us to directly compute the pixel numbers Q_F and Q_M solving the additional problem of **FM quantification**. Moreover the FM quasi-segmentation problem can be solved via definition of a threshold t_{FM} in the grey-scale value histogram segregating X_F as effective as possible from X_M .

For all computations we used the 8-bit CT image version resulting in $N = 256$ possible grey-scale values. Though this is a reduced grey-scale value resolution compared to the original 12-bit version, it yielded a better

filling of the histograms compartments. Since the images were of format 512×512 pixels, we had $Q = 262,144$ pixels bearing grey-scale values for our histograms $X(Y)$.

The signal-to-signal separation measure τ

The classical signal-to-noise ratio (SNR) (see e.g.⁶) in the field of digital image processing is usually defined as the quotient of the mean signal amplitude $\mu(\text{signal})$ and the standard deviation of the noise $\sigma(\text{noise})$:

$$\text{SNR} := \frac{\mu(\text{signal})}{\sigma(\text{noise})} \quad (1)$$

As the SNR was originally designed to quantify the detectability of a narrow signal peak in a broad band of noise in frequency domain, its concept has to be adapted to our detection scenario, where the separability of two distinctly narrow signals peaks has to be measured. We will define something like an analogon τ to the SNR for our FM separation problem. The measure τ will be referred to as a 'signal-to-signal ratio'. The grey-scale value pdfs $f_{\Xi}(y, \Theta)$ of fat ($\Xi = F$) and muscle ($\Xi = M$) in a CT slice are Normal distributions $\Phi(y, E_{\Xi}, S_{\Xi})$ (see⁶) in good approximation, completely determined by their moments E_{Ξ}, S_{Ξ} . Therefore we look for a separability measure τ depending on E_{Ξ}, S_{Ξ} . The separability τ should become better with growing distance of the maxima E_F and E_M . The τ value should be rising with falling widths S_F and S_M making the peaks in the grey value histogram sharper. Both effects lead to smaller overlap of f_F and f_M . The problem shows obvious parallels to the problem addressed by Welch's t-test⁷. Thus the **FM separability** τ , as applied to a CT slice Y , is defined as:

$$\tau(Y) = \frac{E_M - E_F}{\sqrt{\frac{1}{Q_M} S_M^2 + \frac{1}{Q_F} S_F^2}} \quad (2)$$

Inserting the definition of the pixel numbers $Q_F = z \cdot Q$ and $Q_M = (1 - z) \cdot Q$ of the components F and M into (2) we yield:

$$\tau(Y) = \frac{\sqrt{Q} \cdot (E_F - E_M)}{\sqrt{\frac{S_F^2}{z} + \frac{S_M^2}{1-z}}} \quad (3)$$

Since there is no direct access to the actual values of the variables E_F, E_M, S_F, S_M, z one has to approximate them in formula (3) by using estimators $E_F^*, E_M^*, S_F^*, S_M^*, z^*$ instead, yielding an estimator τ^* for τ .

The anatomical FM misclassification measure D

The quality of the quasi-segmentation result $W^*(Y)$ is determined by an informed observer, comparing $W^*(Y)$ to the single CT image Y . This is formalized by the function $d(Y, W^*(Y))$, called the (**single slice FM misclassification**):

$$d(Y, W^*(Y)) := \begin{cases} 0 & \text{if No anatomical error detected comparing } Y \text{ to } W^*(Y) \\ 1 & \text{if Anatomical error detected comparing } Y \text{ to } W^*(Y) \end{cases} \quad (4)$$

Given a CT scan $\underline{Y} = \{Y_1, \dots, Y_L\}$ of an human abdomen we define the (**FM misclassification rate** $D(\underline{Y})$) as the mean of the single-slice FM misclassification:

$$D(\underline{Y}, W^*(\underline{Y})) := \mu_{\underline{Y}}(d(Y, W^*(Y))) = \frac{1}{\#\underline{Y}} \sum_{Y \in \underline{Y}} d(Y, W^*(Y)) = \frac{1}{L} \sum_{Y \in \underline{Y}} d(Y, W^*(Y)) \quad (5)$$

The (**sample FM misclassification rate** $F(\Psi, W^*(\Psi))$) refers to a sample $\Psi = \{\underline{Y}_1, \dots, \underline{Y}_R\}$ of CT scans \underline{Y}_r of different human abdomina analogously:

$$F(\Psi, W^*(\Psi)) := \mu_{\Psi}(D(\underline{Y}, W^*(\underline{Y}))) = \frac{1}{\#\Psi} \sum_{\underline{Y} \in \Psi} D(\underline{Y}, W^*(\underline{Y})) = \frac{1}{R} \sum_{\underline{Y} \in \Psi} D(\underline{Y}, W^*(\underline{Y})) \quad (6)$$

The values $d(Y, W^*(Y))$, $D(\underline{Y}, W^*(\underline{Y}))$ and $F(\Psi, W^*(\Psi))$ depend on the CFES' parameter value α via $W^*(\underline{Y})$ of the scans \underline{Y} in Ψ . If this fact will be emphasized in the following we will append an index α to the respective symbol, as in: $W^*_{\alpha}(\underline{Y})$, $d(Y, W^*_{\alpha}(Y))$, $D(\underline{Y}, W^*_{\alpha}(\underline{Y}))$, $F(\underline{\Psi}, W^*_{\alpha}(\underline{\Psi}))$. If we talk generally about the influence of parameter α on FM misclassification, we use the symbols d_{α} , D_{α} , F_{α} .

The weighted least squares distance S_{α}

The (**WLS distance** $S_{\alpha}(X(Y), X(\Theta))$) of a slice Y with reference to a parameter vector Θ , which was used as target parameter for CFES' nonlinear estimation process, can be thought of as another quality measure defined. It is computed from the empirical histogram $X(Y) := \{x_0(Y), \dots, x_{N-1}(Y)\}$ and the theoretical histogram $X(\Theta) := \{x_0(\Theta), \dots, x_{N-1}(\Theta)\}$ as follows:

$$S_{\alpha}(X(Y), X(\Theta)) := \sum_{n \in G_0} (x_n(Y) - x_n(\Theta))^2 \cdot x_n(Y)^{\alpha} \quad (7)$$

As a seemingly evident hypothesis we primarily assume:

| LFD | Tube-voltage (kVp) | Tube-current (mAs) | Convolution kernel | Spacing between slices (mm) | Slice thickness (mm) |
|-----|--------------------|--------------------|--------------------|-----------------------------|----------------------|
| 1 | 120 | 99 | Standard | 3.00 | 3.75 |
| 2 | 120 | 120 | Standard | 3.00 | 3.75 |
| 3 | 120 | 156 | Standard | 3.75 | 3.75 |
| 5 | 120 | 400 | Standard | 3.00 | 3.75 |
| 6 | 120 | 500 | Standard | 0.625 | 0.625 |
| 7 | 120 | 500 | Standard | 0.625 | 0.625 |
| 8 | 120 | 141 | Bone | 3.00 | 0.625 |
| 9 | 120 | 500 | Standard | 0.625 | 3.75 |
| 10 | 120 | 500 | Standard | 0.625 | 0.625 |
| 11 | 120 | 460 | Standard | 0.625 | 0.625 |
| 12 | 120 | 99 | Standard | 0.625 | 0.625 |
| 13 | 120 | 400 | Standard | 0.625 | 0.625 |
| 14 | 120 | 500 | Bone Plus | 0.625 | 0.625 |
| 15 | 120 | 99 | Standard | 0.625 | 0.625 |
| 17 | 120 | 500 | Standard | 0.625 | 0.625 |
| 18 | 120 | 490 | Standard | 0.625 | 0.625 |
| 19 | 120 | 500 | Standard | 0.625 | 0.625 |
| 20 | 120 | 500 | Standard | 0.625 | 0.625 |
| 21 | 140 | 445 | Bone Plus | 0.625 | 0.625 |
| 22 | 120 | 500 | Standard | 0.625 | 0.625 |
| 23 | 120 | 500 | Standard | 0.625 | 0.625 |
| 24 | 120 | 370 | Standard | 0.625 | 0.625 |
| 26 | 120 | 500 | Standard | 0.625 | 0.625 |
| 27 | 120 | 500 | Standard | 0.625 | 0.625 |
| 28 | 120 | 500 | Standard | 0.625 | 0.625 |
| 29 | 140 | 350 | Standard | 0.625 | 0.625 |
| 30 | 120 | 425 | Standard | 0.625 | 0.625 |
| 31 | 120 | 490 | Standard | 0.625 | 0.625 |
| 32 | 120 | 500 | Standard | 0.625 | 0.625 |
| 33 | 140 | 350 | Standard | 0.625 | 0.625 |
| 34 | 140 | 350 | Standard | 0.625 | 0.625 |
| 35 | 140 | 350 | Standard | 0.625 | 0.625 |

Table 1. CT parameter of the corpse sample (three of the originally 35 corpses excluded).

- (S) For any CT slice Y : A small value of the distance $S_a(X(Y), X(\Theta))$ between measured grey value histogram $X(Y)$ and theoretical grey value histogram $X(\Theta)$ given an actual value of the parameter Θ , or a high value of the FM separability $\tau_a(Y)$, should lead to a good FM quantification in terms of a correctly estimated value of the FM proportion parameter z as well as to a good FM quasi-segmentation measured via the single-slice FM misclassification $d_a(Y, W^*(Y))$.

The minimisation of (7) as well as the formulae for the sensitivity analysis of the WLS approach are described in more detail in the supplementary information. There we derive the equations (S1.3) for small deviations of the WLS parameter estimation, the components of the Jacobi matrix (S2.1)–(S2.13), the components of the estimators covariance matrix (S3.3) and explain our assumption (S4.3) for the covariance matrix of the input, the grey-scale value histogram.

The sample of human body scan data

CFES was applied to CT scans of human abdomens. The abdomen's upper and lower bounds were defined by anatomical landmarks: The central diaphragm region characterizes the upper scan limit, while the visible conjunction of the upper part of os pubis (facies symphysialis) marks the lower limit.

After excluding three of the originally 35 corpses from the sample due to missing CT data or due to artefacts deteriorating the CT scan, the sample consisted of $R = 32$ human bodies of 11 female and 21 male subjects. The prosecution had levied on the bodies and ordered the CT scans as well as the autopsies. For every body age (mean = 57.41 y, stddev. = 20.22 y), body weight (mean = 71.37 kg, stddev. = 16.59 kg), body height (mean = 1.70 m, stddev. = 0.08 m) and body mass index (mean = 24.43 kg/m², stddev. = 4.92 kg/m²) were registered.

For all of the scans \underline{Y} a LightSpeedVCT (GE Medical Systems) was used with a filtered back projection reconstruction, a 'body' filter type and with a spiral pitch factor of 0.984. Table 1 shows an overview of our scan's CT

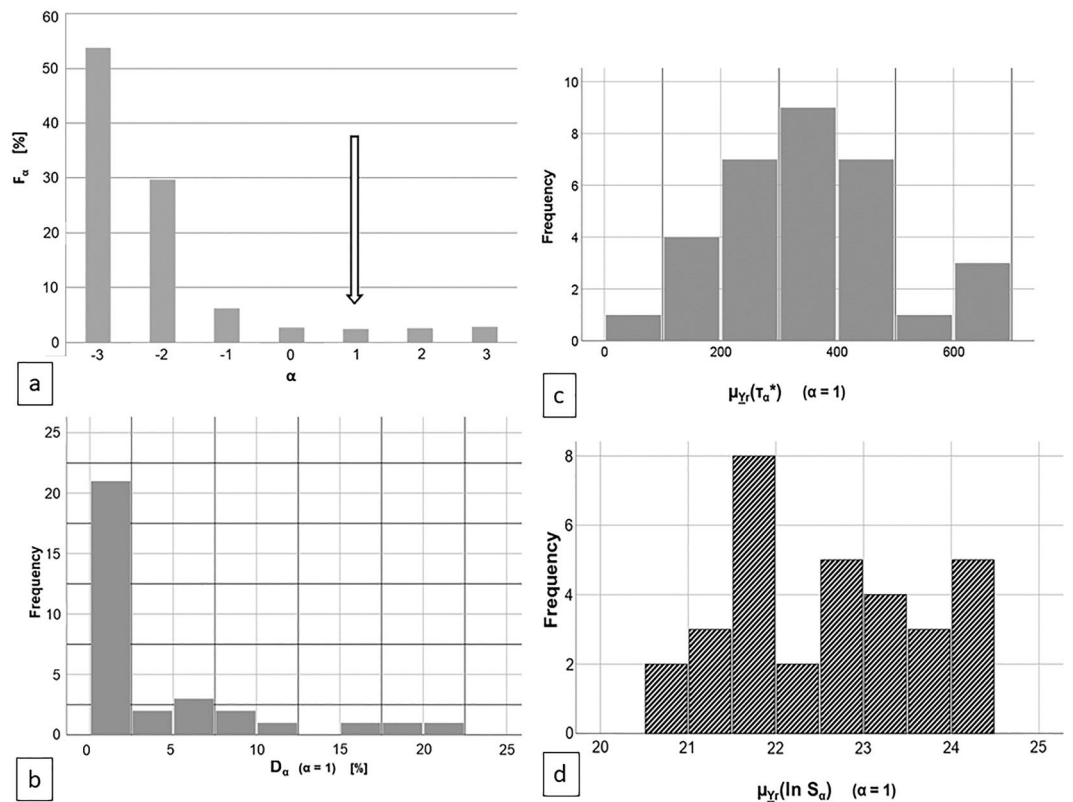


Figure 1. (a) Mean FM misclassification rate F_α for whole body CT sample Ψ as a function of weight exponent α . (b) Histogram of the FM misclassification rate D_α for $\alpha = 1$. (c) Histogram of $\mu_{Y_r}(\tau_\alpha^*)$ with $\alpha = 1$. (d) Histogram of $\mu_{Y_r}(\ln S_\alpha)$ for $\alpha = 1$.

parameter values. For our pixel-volume calculation the parameter ‘Spacing between slices’ (column 5 of Table 1), was relevant rather than the parameter ‘Slice Thickness’ (column 6 of Table 1).

Former usage of study material

The CT-data and its CFES-evaluations were used in the thesis²⁷ of Schenkl, who is the first author of the present article.

Ethics approval

The study was reviewed and approved by the ethics committee of the University Hospital Jena. According to the ethics committee’s statement written consents of the kinship for the abdominal CT slices shown in the article were not needed since the bodies were confiscated by the local prosecution who directed the CT scans for investigations. Moreover, the representations of the slices are totally anonymized.

Results

Quality parameters: single analysis. Figure 1A presents misclassification rate $F_\alpha(\Psi, W^*(\Psi))$ as a function of α . The optimum exponent $\alpha = 1$ with $F_\alpha = 0.036 = 3.6\%$ is marked. Figure 1b is the a misclassification rate D_α histogram for $\alpha = 1$.

Figure 1c, d show the mean separability $\mu_{Y_r}(\tau_\alpha^*)$ - and the mean $\mu_{Y_r}(\ln S_\alpha)$ histogram respectively, taken over all CT slices Y_{r1} for an abdominal scan Y_r for each one of the R bodies.

Quality parameters: multiple parameters analysis

In Fig. 2a a scatterplot of mean separability $\mu_{Y_r}(\tau_\alpha^*)$ (X-axis) versus misclassification rate D_α (Y-axis is shown while Fig. 2b presents a scatterplot of the mean log minimal distance $\mu_{Y_r}(\ln S_\alpha)$ and misclassification rate D_α (for both $\alpha = 1$). Figure 2c is a scatter plot of mean separability $\mu_{Y_r}(\tau_\alpha^*)$ (X-axis) and mean log WLS distance $\mu_{Y_r}(\ln S_\alpha)$ (Y-axis) with $\alpha = 1$ showing a distinctive linear relation (Linear regression analysis: $R^2 = 0.772$, $p < 0.001$, slope $a = 0.007$ and intercept $b = 20.145$).

Figure 2d–f present an example slice with $\mu_{Y_r}(\tau_\alpha^*) = 170.24/D_\alpha = 4\%$ with single-slice misclassification $d_\alpha = 0$ but low mean separability $\tau_\alpha^* = 128.35$ and middle-to-high log distance $\ln S_\alpha = 23.38$ for $\alpha = 1$.

As slope $a > 0$ contradicts the naive intuition (S) (‘high separability τ_α causes low distance S_α ’), we further studied the effect using simulated images.

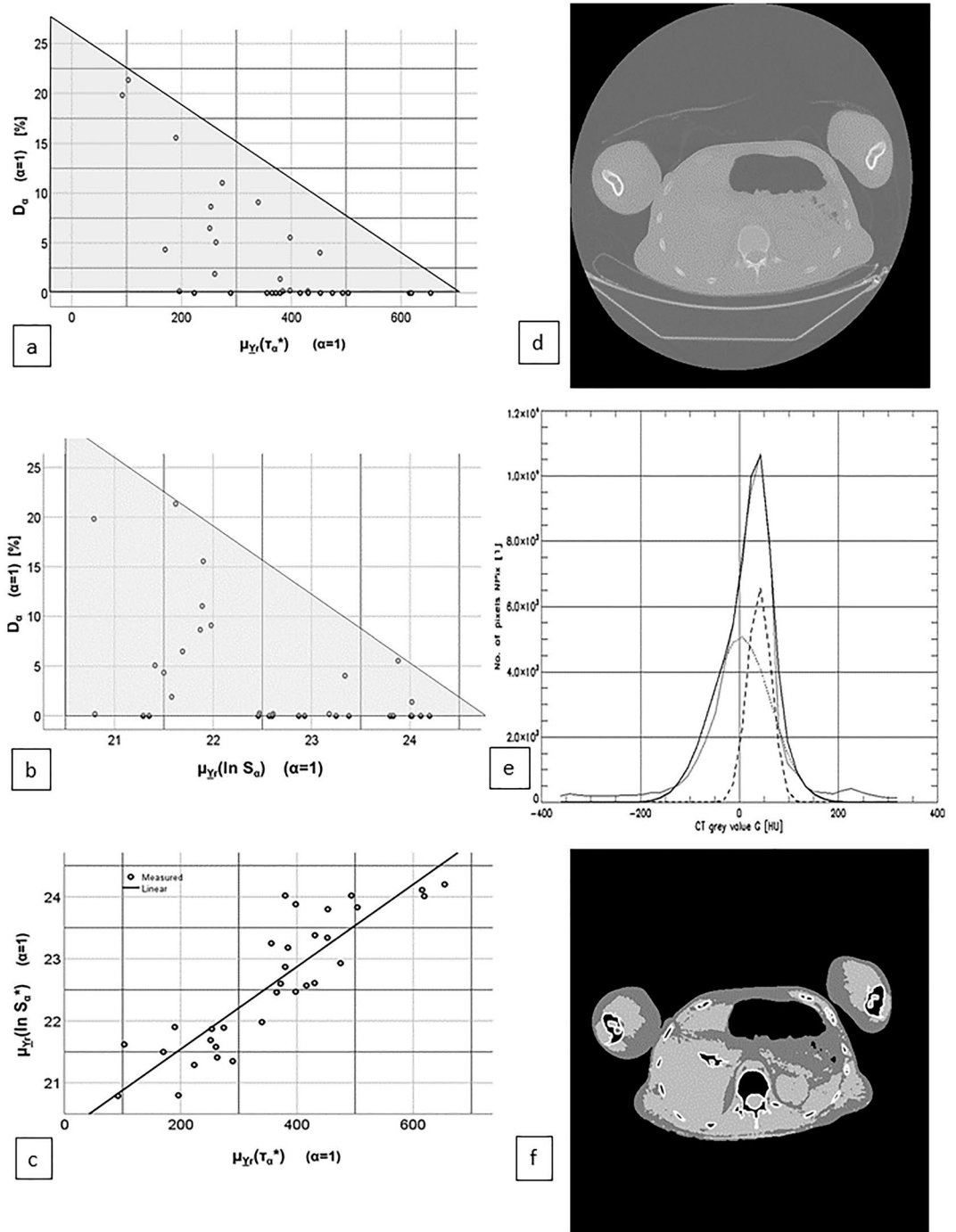


Figure 2. (a) Scatterplot: FM misclassification rate D_α and mean FM separability $\mu_{Y_r}(\tau_\alpha^*)$ of all individual scans Y_r in the sample Ψ for $\alpha=1$. Triangle inserted to show non-function type relation. (b) Scatterplot: FM misclassification rate D_α and mean log WLS distance $\mu_{Y_r}(\ln S_\alpha)$ of all individual scans Y_r in sample Ψ for $\alpha=1$. (c) Scatterplot: mean log WLS distance $\mu_{Y_r}(\ln S_\alpha)$ and mean FM separability $\mu_{Y_r}(\tau_\alpha^*)$ of each individual scan Y_r in the sample Ψ with $\alpha=1$. The drawn line represents a linear model fitted with by a least square approach. (d) Example slice from scan LFD_5: Original CT slice. (e) Example slice from scan LFD_5: Grey value histograms: measured: X (thin line), fitted: X(\ominus) (fat line), partial fat: X_F (\odot) (dotted line), partial muscle X_M (\ominus), fat vertical line: threshold t_{FM} . (f) Example slice from scan LFD_5: Quasi segmented image (light grey: muscle, dark grey: fat, white boundary area: bones), the weight exponent was $\alpha=1$.

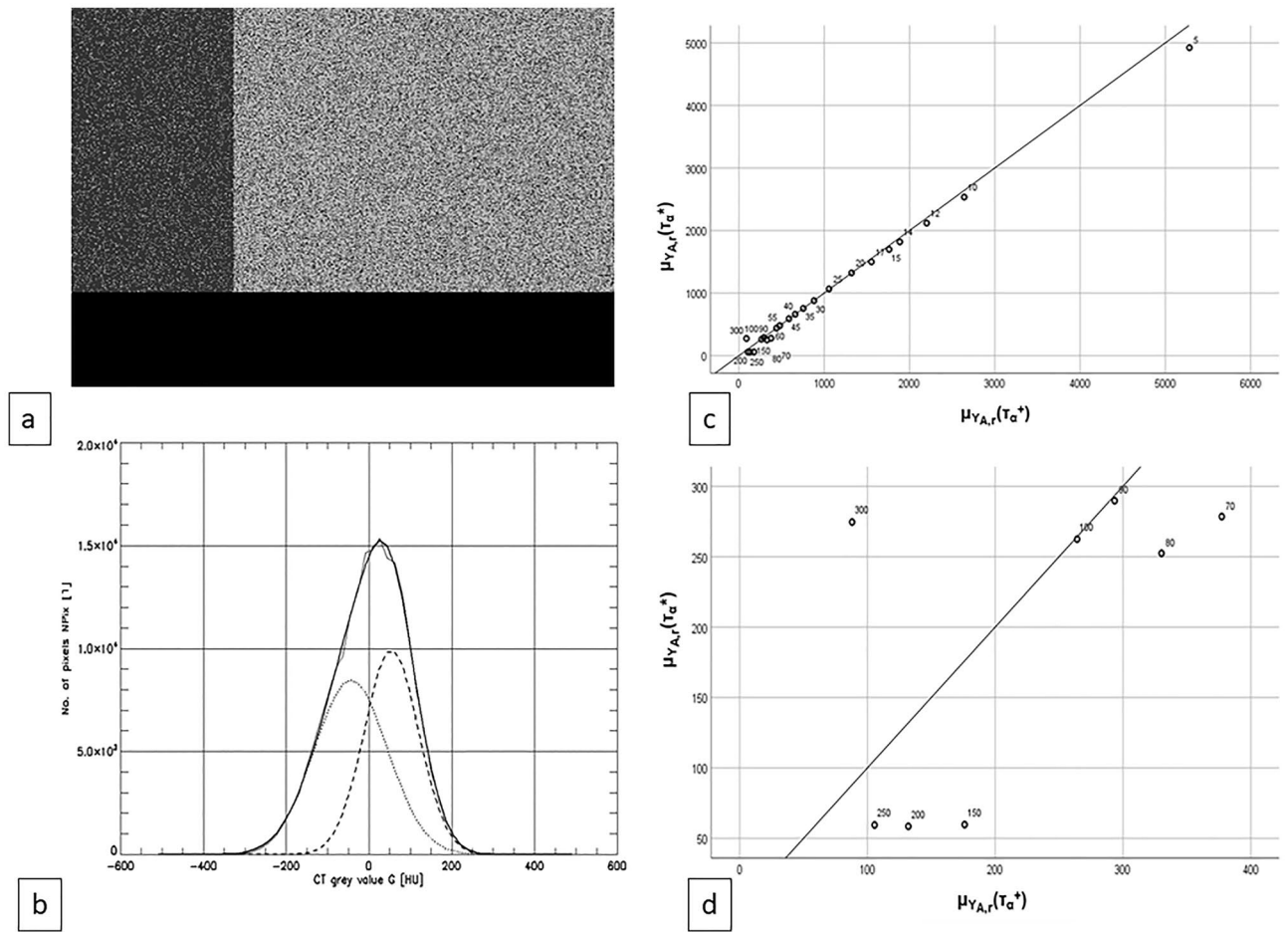


Figure 3. (a) Monte Carlo simulated CT image $Y_{A,r}$ with components fat (dark grey), muscle (light grey), air (black). The true image parameters were: $E_F^+ = -80$ HU, $E_M^+ = 50$ HU, $S_F^+ = S_M^+ = 70$ HU, $z^+ = 0.3$. (b) Grey-scale value histograms X (empirical), $X(\Theta)$ (fitted via CFES with $\alpha = 1$), $X_F(\Theta)$, $X_M(\Theta)$ with: Thin line: measured, fat line: estimated, dotted: fat, dashed: muscle. (c and d) Scatterplots of mean true value $\mu_{Y_{A,r}}(\tau_\alpha^+)$ vs. mean estimated value $\mu_{Y_{A,r}}(\tau_\alpha^*)$ with $z = 0.3$. (c) The diagrams containing all of the synthetic images $Y_{A,r}$ generated. (d) Diagram as in the upper right part only diagrams $Y_{A,r}$ were considered with standard deviations $S = S_F^+ = S_M^+ \geq 70$ HU. The parameter $S = S_F^+ = S_M^+$ labels the data-points. The weight exponent $\alpha = 1$ was used in S_α .

Quality parameters: Monte Carlo analysis of mean log WLS distance $\mu_{Y_r}(\ln S_\alpha)$ and mean FM separability $\mu_{Y_r}(\tau_\alpha^*)$

We used two series $Y_{A,1}, \dots, Y_{A,R}$ and $Y_{B,1}, \dots, Y_{B,R}$ ($R = 20$) of simulated images (Fig. 3a,b) with $E_F = -80$ HU, $E_M = 50$ HU and $S = S_F = S_M = 10$ HU, 12 HU, 14 HU, 15 HU, 17 HU, 20 HU, 25 HU, 30 HU, 35 HU, 40 HU, 45 HU, 55 HU, 60 HU, 70 HU, 80 HU, 90 HU, 100 HU, 150 HU, 200 HU, 250 HU. For series A and B respectively z was set $z = 0.3$ and $z = 0.5$. All images $Y_{A,r}$ and $Y_{B,r}$ were evaluated by CFES algorithm with $\alpha = 1$. The true values and the estimated values of E_F , E_M , S_F , S_M , z , and the quality parameters $\ln S_\alpha$, the true and the estimated value of τ_α were recorded.

The τ_α^+ vs. τ_α^* —scatterplots in Fig. 3c, d respectively represent the full sample $Y_{A,r}$ ($r = 1, \dots, R$) and the $Y_{A,r}$ with τ_α^+ -range of $[0, 400]$ only.

Figure 4a is the scatterplot of $\ln S_\alpha$ vs. τ_α over the $Y_{A,r}$ —and the $Y_{B,r}$ sample (A: $z = 0.3$: fat rings, B: $z = 0.5$: fat rectangles). The standard deviation value S labels every data point. A dashed rectangle marks the area $\times [0, 700] \times [20.5, 24.5]$ of Fig. 2c.

The center of Fig. 4b repeats Fig. 4a with associated histograms X (thin line), $X_\Xi^*(\Theta)$ (dotted: $\Xi = F$, dashed: $\Xi = M$), $X^*(\Theta)$ shown for the marks $S = 12$ HU, 25 HU, 40 HU, 60 HU, 100 HU.

Figure 4c investigates the ‘umbrella’ in Fig. 4a, b. It shows a 3-D scatterplot ($Y_{A,r}$ or $Y_{B,r}$, excluded if $S \leq 5$ or $S \geq 100$ or if $RE_z \geq 0.2$) of separability τ_α^* , logarithmic distance $\ln S_\alpha$ and the relative error $RE_z = |z^+ - z^*|/z^+$ of z -estimation (Black filled bubbles = $Y_{A,r}$ ($z = 0.3$), void bubbles = $Y_{B,r}$ ($z = 0.5$)). The view is leading through the left and right front sides as well as through the top side of the cube.

Figure 4d presents the Gaussian estimation of our WLS estimation z^* 's standard deviation $S(z^*)$ with $\alpha = 0$, where WLS = LS and $E_F = -80$ HU; $E_M = 50$ HU, $z = 0.3$ (graph for $z = 0.5$ is nearly identical), $S = S_F = S_M$ varied between $S = 20$ HU and $S = 90$ HU (Fig. 4d left) and with $\Delta E := E_M - E_F$ between $S = 20$ HU and $S = \Delta E/2 = 130$ HU

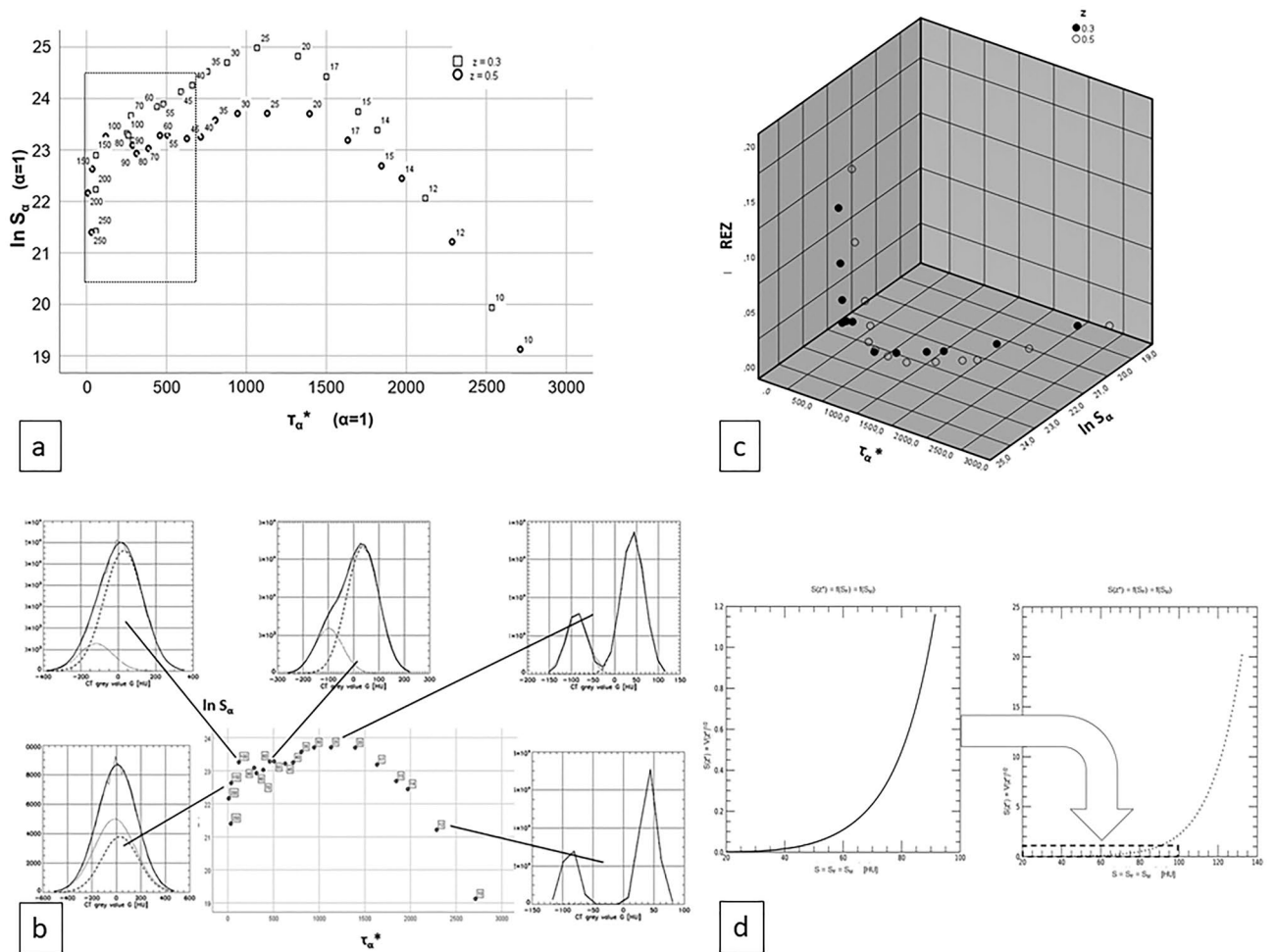


Figure 4. (a) Scatterplot for the variables FM separability τ_{α}^* and logarithm of weighted least squares sum $\ln S_{\alpha}$ of the Monte Carlo images $Y_{A,r}$ and $Y_{B,r}$ with $r = 1, \dots, 20$ (where A: $z = 0.3$ and B: $z = 0.5$). The data-points are labelled with the number of their grey-scale value histograms true standard deviations $S^+ = S_F^+ = S_M^+$. The dashed rectangle on the left side marks the $\tau_{\alpha}^* - \ln S_{\alpha}$ -area of the diagram in Fig. 2d. (b) In the middle: Scatterplot for the variables FM separability τ_{α}^* and logarithm of weighted least squares sum $\ln S_{\alpha}$ of the Monte Carlo images $Y_{A,r}$ with $r = 1, \dots, 20$. The data-dots are labelled with the value of their grey-scale value histograms true standard deviations $S^+ = S_F^+ = S_M^+$. Encircling the central scatterplot five empirical and estimated grey-scale value histograms X (thin, drawn), $X(\Theta)$ (fat, drawn), $X_F(\Theta)$ (fat, dashed), $X_M(\Theta)$ (fat, dotted) are presented, each is connected to its corresponding data-dot in the central scatterplot by a fat drawn pointer-line. (c) 3-D scatterplot of the variables FM separability estimator τ_{α}^* , the logarithm of the weighted least squares distance $\ln S_{\alpha}$ and relative error RE_z of the z -estimator for the image-series $Y_{A,r}$ and $Y_{B,r}$ with $r = 1, \dots, R$. Black bubbles: A: $z = 0.3$ void bubbles: B: $z = 0.5$. Images $Y_{A,r}$ and $Y_{B,r}$ were excluded if $S \leq 5$ or $S \geq 100$ as well as if $RE_z \geq 0.2$. The weighting exponent was set to $\alpha = 1$. (d) Computed diagrams of the z -estimator's standard deviation $S(z^*)$ as a function of the standard deviations $S = S_F = S_M$ of the partial fat and muscle distributions in the scenario Y_A and Y_B with $E_F = -80$ HU, $E_M = 50$ HU, $z = 0$ and ranges $20 \text{ HU} \leq S \leq 90 \text{ HU}$ (A: left) and $20 \text{ HU} \leq S \leq 130 \text{ HU}$ (B: right). Note the different ranges on the ordinate. The graphs were generated with a true fat ratio value $z^+ = 0.3$. However the choice of the true fat ratio value z^+ had virtually no influence on the result when it was changed from $z^+ = 0.3$ to $z^+ = 0.5$. The weight exponent was chosen $\alpha = 0$ which means z^* is an LS-estimator rather than an WLS estimator as it is in the case $\alpha = 1$.

respectively (Fig. 4d right). The curves in Fig. 4d were generated for $\alpha = 0$, since this guarantees the weighting to be of no influence. The latter was preferable as for each of the values $r = 1-3$ the weight-matrix else would have been different for every two slices else, thus introducing noise in the evaluation.

General results

As examples of the CFES performance Fig. 5a,d,g,j show four slices Y , their grey-scale value histograms $X(Y)$, $X(\Theta^*)$, $X_F(\Theta^*)$, $X_M(\Theta^*)$ (Fig. 5b,e,h,k) and the respective quasi-segmentation results $W^*(Y)$ (Fig. 5c,f,i,l). Horizontal image-line 1, 2, 3, 4 corresponds to LFD 06, 22, 24, 16 (see Table 1) respectively. While the first three slices (a), (d), (g) lead to acceptable results, slice (j) was added as a poor result caused by weak contrast and artefacts. LFD 16 was excluded from our statistical evaluations.

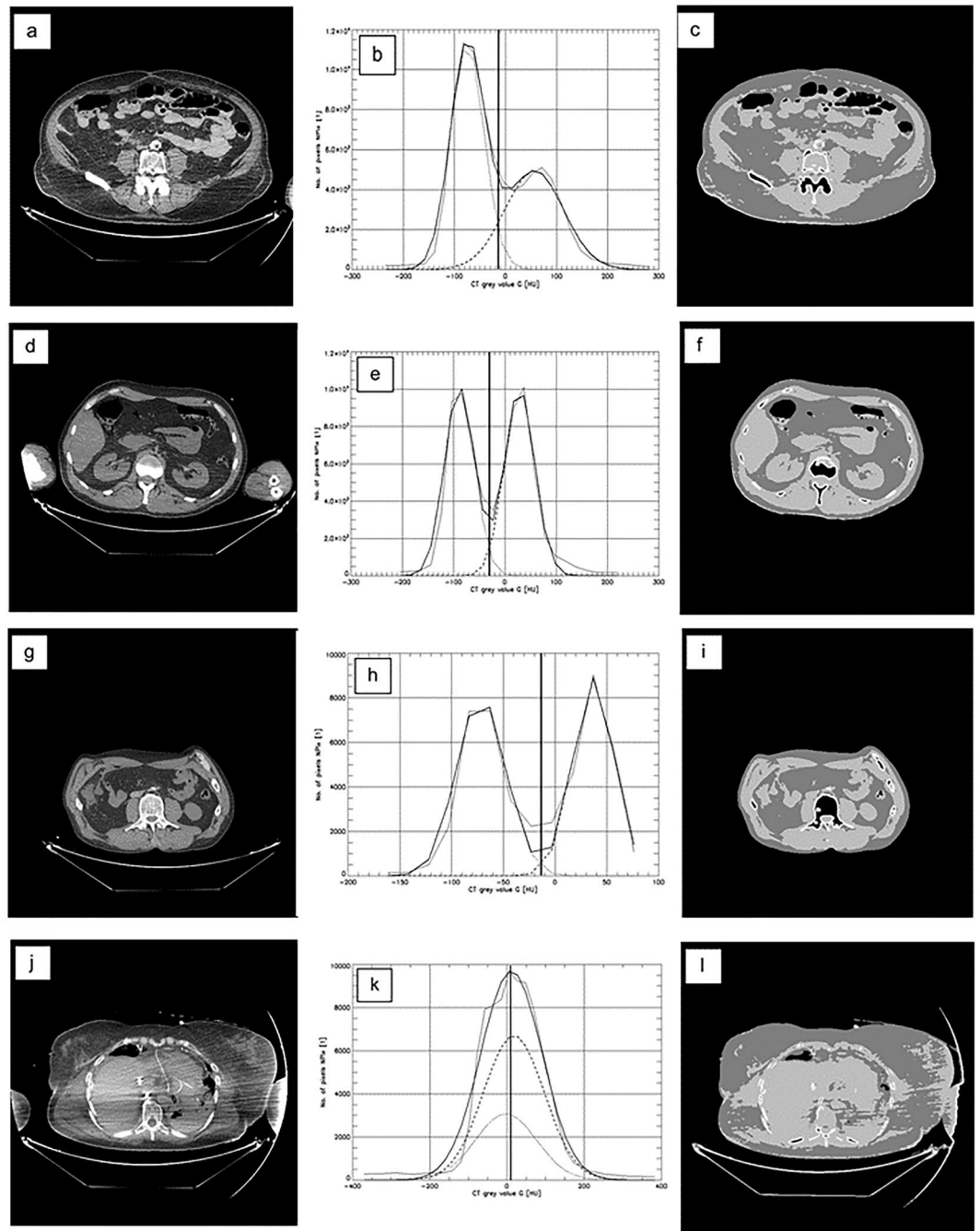


Figure 5. Four slices Y in (a), (d), (g), (j) from four different persons at different abdominal positions. Respective histograms with thin line = measured data X , thick line = fitted histogram $X(\Theta^*)$, partial histograms $X_F(\Theta^*)$, $X_M(\Theta^*)$ in (b), (e), (h), (k) and quasi-segmentation results $W^*(Y)$ in (c), (f), (i), (l). The parameter α was $\alpha = 1$.

Discussion

General discussion. The exponent α 's calibration in the paper⁴ yielded optimum values $\alpha = -1$ for synthetic CT images and $\alpha = -1$ for the real-world phantom. Application of CFES to our body CT scan sample Ψ and evaluating the misclassification rate $F_\alpha(\Psi, W^*(\Psi))$ leads to Fig. 1a. This indicates $\alpha = 1$ to be the optimal parameter value with $F_\alpha = 3.6\%$ for our body sample though the values for $\alpha = -1$, $\alpha = 2$ and $\alpha = 3$ lie only slightly above: The F_α -values for $\alpha = -3, -2, -1$ are $F_\alpha = 54\%, 30\%, 6\%$ which is distinctively higher.

The three quality measures τ_α , D_α and S_α are a posteriori quantities. This is evident for D_α , and for S_α . Only for synthetic CT images the separability τ_α^+ could be determined in advance. τ_α can only be estimated from a priori parameter estimators as described in⁴. The parameter τ_α obtains its justification from analogies in signal

| Author, year, citation | Quality measure |
|--|--|
| Otsu ⁸ , 1979 | Variance analysis inspired, threshold k optimization of $\eta(k)$, $\eta(k) := V_{\text{interclass}}/V_{\text{Total of levels}}$ non-parametric |
| Mitsopoulos ⁹ , 1998 | Comparison to sliced cadavers, correlation analysis |
| Rogalla ¹⁰ , 1998 | Comparison to phantom, standard error |
| Yoshizumi ¹¹ , 1999 | Comparison to cadavers, planimetry, correlation analysis |
| Glasbey ¹² , 2002 | Comparison to simulated CT images, RMSE, bias, std.deviation |
| Dobrowolsky ¹³ , 2004; Romvary ¹⁴ , 2006 | Comparison to butchered pigs, fat-muscle ratio, linear regression |
| Pednekar ¹⁵ , 2005 | Comparison to manual segmented images, true negative/positive pixels |
| Johansen ¹⁶ , 2007 | Comparison to estimation on carcasses, linear regression |
| Kongsro ¹⁷ , 2008 | Comparison to measurement results of manual dissections, linear regression |
| McEvoy ¹⁸ , 2008 | Std. dev. of local maximum estimator in grey-scale value histogram |
| Oshima ¹⁹ , 2008 | No measure published |
| McEvoy ²⁰ , 2009 | No measure published |
| Checchini ²¹ , 2011 | Correlation analysis estimated fat volume vs. BMI |
| Kanaly ²² , 2011 | Comparison to informed observers results |
| Subramaniam ²³ , 2012 | Lung tissue inhomogeneity measurement method, validated by distinguishability of end-inspiration- and end-expiration heterogeneity |
| Kim ²⁴ , 2013 | Comparison to dual-energy X-ray absorptiometry results, correlation analysis |
| Kullberg ²⁵ , 2017 | Comparison to manual CT-segmentation results, correlation analysis |

Table 2. Studies in literature about CT fat segmentation: quality measures.

processing and in statistics in contrast to the WLS difference S_α , whose minimality is necessary for estimating the parameter vector Θ . The misclassification rate D_α is relying on a priori expert knowledge.

Quality measures should quantify firstly the suitability of the input data (the separability of fat from muscle in a CT-image), secondly the efficacy of the segmentation (WLS fitting of the model $X(\Theta)$ to the histogram X and computation of the threshold t_{FM}) and at last the correctness of the result (the anatomical correctness of FM segmentation). Our measures τ_α , S_α and D_α cover all three aspects: τ_α measures the shape of the $X(Y)$ to quantify the CT-image Y 's suitability for CFES application, S_α judges the final 'distance' between X and $X(\Theta)$ only, while D_α evaluates $W^*(Y)$ without distinction of genesis.

In⁴ we gave a synopsis of tissue segmentation studies which we review in Table 2 for their quality measures.

Our choice of τ_α , S_α and D_α may be partly explained since we were not able to determine a gold standard as e.g. manual dissection. S_α as the minimized quantity of quasi-segmentation seems to be a natural choice. As we knew the single material grey-scale value pdf to be Gaussian, we did not want to use non-parametric fitting methods as e.g.⁸. Usage of τ_α was inspired by the t-test statistic, used for deciding whether two Gaussians have different expectation values. Moreover there are similarities of τ_α to the signal theoretic quantity of signal to noise ratio (SNR), quantifying the separability of two signals in frequency domain. Finally the misclassification rate D_α seemed to be a canonical choice for a measure of the pure CFES-output quality.

In the slice of Fig. 2d–f the d_α implies good quality, whereas S_α indicates low quality, therefore we will investigate in more detail the interrelations between the three quality measures. Yet for α -calibration there is no better parameter than D_α or F_α since the usage of S_α as a quality measure for fat quantification runs into problems we will discuss below.

As CT numbers generally depend on temperature (see²⁶), the question arises whether the results presented are robust against sample temperature changes. The results in²⁶, using conventional polychromatic 120kVp CT (CPI) (as well as virtual monoenergetic images by dual energy measurements (VMI) with 40kVp, 50 kVp, 60 kVp, 70 kVp, 100 kVp, 120 kVp, 140 kVp), seem to give evidence to the hypothesis of robustness: The temperature-changes in an interval of [10 °C, 60 °C] led to changes of [–40 HU, –80 HU] in the grey-level values for fat-substitute and of [50 HU, 60 HU] for muscle-substitute in CPI (see Fig. 5 in²⁶) which lie in the range of inter-personal grey-level value variations of our study. Furthermore²⁶ indicates that there is no influence of the convolution kernel chosen on the temperature dependence of the grey-level values (see Fig. 5 in²⁶) for CPI. The paper²⁶ implies possibilities for further research on quasi-segmentation e.g. by using VMI and/or temperature control for improving CFES-results.

Relations between the three quality measures

In Fig. 2a the sample's location in the triangle tells us that high values of the $\mu_{Yr}(\tau_\alpha^*)$ seem to force low values on D_α (see (8A)), whereas high values of D_α constrain $\mu_{Yr}(\tau_\alpha^*)$ to low values (see (8B)). Nevertheless, there are abdominae with low values of D_α and of $\mu_{Yr}(\tau_\alpha^*)$ (see (8C)):

$$\mu_{Yr}(\tau_\alpha^*) \uparrow \Rightarrow D_\alpha \downarrow \quad (8A)$$

$$D_\alpha \uparrow \Rightarrow \mu_{Yr}(\tau_\alpha^*) \downarrow \quad (8B)$$

$$\exists : D_\alpha \downarrow \wedge \mu_{Y_r}(\tau_\alpha^*) \downarrow \tag{8C}$$

While (8A) and (8B) match our intuition (S) about relations of the quality measures, the last vertex of the triangle [(8C): low $\mu_{Y_r}(\tau_\alpha^*)$ and low D_α] seems to contradict them grossly at first sight. Note however, that the minimum value of $\mu_{Y_r}(\tau_\alpha^*)$ for $D_\alpha = 0\%$ is $\mu_{Y_r}(\tau_\alpha^*) \approx 200$. A ‘linear approximation’ to the upper bound of the relation is for $\alpha = 1$:

$$0 \leq D_\alpha \leq A \cdot \mu_{Y_r}(\tau_\alpha^*) + B \tag{9}$$

where the line $A \mu_{Y_r}(\tau_\alpha^*) + B$ is estimated by connecting the two data-points $(x_0, y_0) = (103.0, 21.35\%)$ for scan LFD_19 and $(x_1, y_1) = (654.0, 0.0\%)$ for scan LFD_23 by:

$$A^* = \frac{y_1 - y_0}{x_1 - x_0} = \frac{21.35\% - 0.0\%}{103.0 - 654.0} = -0.03875\% \tag{10}$$

$$B^* = y_1 - A^* \cdot x_1 = 0.0 + 0.03875 \cdot 654.0 = 25.3425\% \tag{11}$$

Figure 2d–f displays slice 119 of scan LFD_5 as an example of a slice $Y_{r,i}$ with low d_α and with low τ_α^* as well. This suggests the hypothesis of the threshold t_{FM} to be robust against fluctuations of τ_α .

Figure 2b is a scatterplot of $\mu_{Y_r}(\ln S_\alpha)$ vs. D_α for $\alpha = 1$. As in Fig. 2a we see a relation, not a functional graph but a ‘formed cloud’ filling one corner of the first quadrant. In analogy to Fig. 2a we see on Fig. 2b high values of $\mu_{Y_r}(\ln S_\alpha)$ seem to force low values of D_α [see (12A)], contradicting intuition (S), whereas high values of D_α constrain low values of $\mu_{Y_r}(\ln S_\alpha)$ [see (12B)] contradicting (S) as well. The third corner of the triangle meets intuition (S): low $\mu_{Y_r}(\ln S_\alpha)$ coincides with low D_α [see (12C)].

$$\mu_{Y_r}(\ln S_\alpha) \uparrow \Rightarrow D_\alpha \downarrow \tag{12A}$$

$$D_\alpha \uparrow \Rightarrow \mu_{Y_r}(\ln S_\alpha) \downarrow \tag{12B}$$

$$\exists : \mu_{Y_r}(\ln S_\alpha) \downarrow \wedge D_\alpha \downarrow \tag{12C}$$

The upper boundary approximation line of the cloud is:

$$0 \leq D_\alpha \leq C \cdot \mu_{Y_r}(\ln S_\alpha) + G \tag{13}$$

Estimating the line $C \mu_{Y_r}(\ln S_\alpha) + G$ by connecting the data-points $(x_0, y_0) = (21.62, 21.35\%)$ (scan LFD_19) and $(x_1, y_1) = (23.88, 5.6\%)$ (scan LFD_27) gives:

$$C^* = \frac{y_1 - y_0}{x_1 - x_0} = \frac{5.6\% - 21.35\%}{23.88 - 21.62} = -6.97\% \tag{14}$$

$$G^* = y_1 - C^* \cdot x_1 = 5.6\% + 6.97\% \cdot 23.88 = 172.0\% \tag{15}$$

Figure 2c is the scatterplot ($\alpha = 1$) of $\mu_{Y_r}(\ln S_\alpha)$ vs. $\mu_{Y_r}(\tau_\alpha^*)$. The distinct linear relationship with positive slope was significant in by linear regression analysis, which is a contradiction (16) to (S):

$$\mu_{Y_r}(\tau_\alpha^*) \uparrow \Rightarrow \mu_{Y_r}(\ln S_\alpha) \uparrow \tag{16}$$

The deviations (8C), (12A), (12B), (16) from (S) were investigated, using the simulated CT images $(Y_{A,r})_{r=1,\dots,R}$ and $(Y_{B,r})_{r=1,\dots,R}$ (see Fig. 3a, b).

Figure 3c shows a scatterplot of τ_α^+ vs. τ_α^* with ranges $[0, 6000]$ on the X- and on the Y-axis for sample $Y_{A,r}$. The matching of τ_α^+ and τ_α^* is almost perfect here, becoming slightly weaker in the region near (0,0). Zooming in on a range of $[0, 400]$ for τ_α^+ we notice strong deviations of the estimator values τ_α^* in Fig. 3d. Note that the range of our sample Ψ ’s mean separabilities $\mu_{Y_r}(\tau_\alpha^*)$ is $[0, 600]$: The correlation of τ_α^+ and τ_α^* of our sample Ψ would have to be estimated from the zoomed-in representation in Fig. 3d rather than from the diagram in Fig. 3c. The correlation of τ_α^+ and τ_α^* refutes the objection to the positive correlation $\mu_{Y_r}(\tau_\alpha^+) - \mu_{Y_r}(\ln S_\alpha)$ in Fig. 2c: “Maybe there is a negative correlation in the analogous $\mu_{Y_r}(\tau_\alpha^+) - \mu_{Y_r}(\ln S_\alpha)$ diagram, which cannot be detected as τ_α^+ is not accessible for our real-world-sample Ψ , but the $\mu_{Y_r}(\tau_\alpha^+) - \mu_{Y_r}(\ln S_\alpha)$ diagram shows a positive correlation caused by a flaw of CFES generating a locally negative correlation between τ_α^+ and τ_α^* ?” Despite the noise in Fig. 2c there is a distinctly positive relation, so the objection can be rejected.

The relations between τ_α^* and $\ln S_\alpha$ are demonstrated in Fig. 4a as scatterplot of τ_α^* vs. $\ln S_\alpha$ with $\alpha = 1$ for the images $Y_{A,r}$ (square marks). The partial graph over τ_α^* -range $[1150, 3000]$ corresponds with (S): Rising τ_α^* is correlated with falling $\ln S_\alpha$ and vice versa:

$$\forall \tau_\alpha^* \in [1150, 3000] : (\tau_\alpha^* \uparrow \Leftrightarrow \ln S_\alpha \downarrow) \tag{17}$$

The curve in the range of $[1150, 3000]$ is a function graph without random noise.

In the complementary partial graph over τ_α^* -range $[0, 1150]$ the curve is more noisy. Its tendency contradicts (S) since rising τ_α^* coincides with rising $\ln S_\alpha$:

$$\forall \tau_{\alpha}^* \in [0, 1150] : (\tau_{\alpha}^* \uparrow \Leftrightarrow \ln S_{\alpha} \uparrow) \quad (18)$$

The two τ_{α}^* -ranges are associated to two ranges of $S = S_F = S_M$ in the synthetic sample A: ' τ_{α}^* in [1150, 3000]' corresponds to ' S in [10 HU, 25 HU] whereas ' τ_{α}^* in [0, 1150]' is associated to $S \geq 25$ HU. Though $\Delta E = E_M - E_F = E_F^+ = 50$ HU $-(-80$ HU) = 130 HU, even $S = 25$ HU means only a moderate grey-scale value distributions densities f_M and f_F overlap.

The area of the sample Ψ 's scatterplot $\mu_{Y_r}(\tau_{\alpha}^*)$ vs. $\mu_{Y_r}(\ln S_{\alpha})$ in Fig. 2c can be realized as the dashed rectangle in the τ_{α}^* vs. $\ln S_{\alpha}$ —scatterplot Fig. 4a: This area lies on the left side of [0, 1150] with trend (17) and comprises a vast variation of the $\ln S_{\alpha}$ —values in the scatterplot of the synthetic samples A and B. The trend in Fig. 2c seems to be fairly linear despite varying fat-proportions z in the real-worlds CT scan sample Ψ . It does not depend on the fat-proportion z , since the umbrella-form with the fuzzy left end shows up in both plots of Fig. 4a. While the top of the umbrella at about $\tau_{\alpha}^* = 25$ sank from ca. $\ln S_{\alpha} = 25$ to ca. $\ln S_{\alpha} = 23.7$ for z rising from $z = 0.3$ to 0.5 , the maximum-location on the τ_{α}^* -axis is nearly identical.

To see, how the overlap between $z \cdot f_F$ and $(1 - z) \cdot f_M$ of a slice is associated to its position on the umbrella, Fig. 4b shows the τ_{α}^* vs. $\ln S_{\alpha}$ scatterplot of $(Y_{A,r})_{r=0,\dots,R}$ again. Additionally, we present five of the grey-scale value histograms $X(Y)$, $X(\Theta)$, $X_F(\Theta)$, $X_M(\Theta)$ associated to the scatterplots data-points for the parameter values $S = S_F = S_M = 12$ HU, 25 HU, 60 HU, 100 HU, 150 HU. The umbrella's top lies at $S = 25$ HU, and it is obvious, that the *increasing overlap* of the partial histograms $X_F(\Theta)$ and $X_M(\Theta)$ increasingly deteriorates z -estimation quality when moving from the right to the left side of the umbrella. The z -estimation quality can be recognized roughly from the overlap area under $X_F(\Theta)$ and $X_M(\Theta)$.

The importance of the umbrella—phenomenon can be demonstrated by relating it to the relative error $RE_z := |z^* - z|/z^*$ of our target variable z . In Fig. 4c, a 3-D scatterplot of τ_{α}^* vs. $\ln S_{\alpha}$ vs. RE_z for $(Y_{A,r})_{r=1,\dots,R}$ (black bubbles) and $(Y_{A,r})_{r=1,\dots,R}$ (white bubbles) with the additional constraints: $5 \text{ HU} \leq S \leq 100 \text{ HU}$ and $RE_z < 0.2$ is shown. It demonstrates the abrupt strong rising of the error RE_z in the target variable z^* if the data points leave the right side of the 'umbrella' in Fig. 4a (which can be found projected on the bottom of the coordinate cube in Fig. 4c as well) and proceed to the left side.

As reason for the umbrella form of the τ_{α}^* vs. $\ln S_{\alpha}$ scatterplot (Fig. 4a), for the fuzziness and for the rise of the relative error RE_z on the left side of the umbrella (in Fig. 4c), a sort of *pseudo-overfitting* of the model $f(y, \Theta)$ can be hypothesized. While the well-known effect of overfitting is associated to using a model with too many parameters, our self-coined term 'pseudo-overfitting' means usage of a model with the correct number of parameters (in our case: 5 parameters: E_F , E_M , S_F , S_M , z) if the model is a mixture of distributions of the same type (in our case: the normal distribution mixture $f(g, \Theta) = z f_F(g, \Theta) + (1 - z) f_M(g, \Theta)$) and with an empirical distribution geometrically similar to a (mixture of less) distribution(s) of the same type (in our case: $X(Y)$ seems to consist of only one normal distribution like in Fig. 4b the grey value distribution for $S = 150$ HU). Here we have to take into account similar effects as those caused by classical overfitting: The fitting process uses the seemingly redundant parameters to model some of the apparently random deviations, which leads to virtually getting 'better' results in terms of S_{α} though the separability τ_{α} decreases (see Figs. 2c and 4a). Associated with smaller values of S_{α} and τ_{α} one notes a dramatic loss in the estimators z^* 's exactness which was measured by the relative error RE_z of the z -estimator z^* (see Fig. 4d).

A principal constraint in fitting Gaussian mixtures

A large proportion of the problems reported in 4.1 and 4.2 seems to come from a principal difficulty in fitting Gaussian mixtures as we did for estimating the parameter vector Θ : The grey value histogram $X(Y)$ of a CT slice Y containing fat- and muscle components only was assumed to be a mixture of the Gaussian components X_F and X_M for fat F and muscle M respectively. Figure 4d demonstrates an exponential growth of the z -estimator's standard deviation $S(z^*)$ for linearly increasing standard deviation $S = S_F = S_M$ of the F - and the of the M -component. This is made plausible by comparing the graph on the left side of Fig. 4 (lower right part) ($z = 0.5$, $20 \text{ HU} \leq S \leq 90 \text{ HU}$) to the graph on the right side ($z = 0.5$, $20 \text{ HU} \leq S \leq 130 \text{ HU}$): though the curve seems to be geometrically identical, the ordinates range has grown by a factor of order 20 while the domain was enlarged from [20 HU, 90 HU] to [20 HU, 130 HU] only. This effect does not depend on the actual value of z : The shown graphs with a true parameter value of $z = 0.3$ are virtually indistinguishable from analogous ones—which therefore are not shown here—with true parameter value $z = 0.5$. The computation results presented in Fig. 4d strikingly demonstrate the fact that the exactness $S(z^*)$ of z -estimation is rapidly deteriorating with increasing overlap of the single-material—normal distributions in the grey-scale value histogram. Given a distance of $\Delta E = E_M - E_F = 50 \text{ HU} - (-80 \text{ HU}) = 130 \text{ HU}$ between the maxima in the histogram, a value of $S = S_F = S_M \geq 60 \text{ HU}$ leads to an approximate z -estimator standard deviation $S(z^*) \geq 0.1$ while $S \geq 90 \text{ HU}$ causes $S(z^*) \geq 1.0$. Note that z is a ratio from the interval [0, 1] which makes an estimator z^* with $S(z^*) \geq 1$ nearly useless.

The difficulty in fitting Gaussian mixtures discussed above and particularly the swiftly growing standard deviation $S(z^*)$ of the z -estimator with rising standard deviation of $S_F = S_M$ seems to be no consequence of our special fitting algorithm WLS chosen: Due to the high number of pixels in a CT slice it is possible to apply asymptotic propositions to the probability distribution of the estimator. Jennrich²⁸ shows, that if the data's (in our case the grey-value histogram X) probability distribution lies in the exponential family (the multinomial distribution of our case belongs to this family) the Fisher Scoring algorithm for iteratively finding the Maximum Likelihood Estimator (MLE) value is identical to the Gauss–Newton algorithm which iteratively detects the Non-linear Least Squares estimator's (NLLS) value, which is detected by CFES as well. Hence the probability distribution of z^* computed here via CFES (NLLS estimation respectively) is nearly identical to the distribution of the MLE of z which *asymptotically is the most effective estimator* (see^{29,30}), at least in case $\alpha = 0$.

Note further that the standard deviation $S(z^*)$'s rapid rising with growing overlap between H_F and H_M are not reported in the fat segmentation literature (see^{8–25}) except maybe hinted in⁸.

A practical consequence of this issue will be to choose CT parameter values minimizing the single-material distribution standard deviations S_F and S_M .

Data availability

The datasets analysed during the current study are not publicly available since they consist of individual CT-scans. Each single slice is linked to a DICOM-file containing experiment parameter data and sensible case information. The DICOM-files are essential for the computations. The bodies were confiscated by the local prosecution who directed the CT scans for investigations.

Code availability

The program code is available from the corresponding author on reasonable request.

Received: 18 March 2022; Accepted: 14 November 2022

Published online: 23 November 2022

References

- Hubig, M., Muggenthaler, H., and Mall, G. Chapter 6.2: Finite element method in temperature-based death time estimation, In: *Estimation of the Time Since Death* (ed. Madea, B.) 114–133, 3rd edn. (CRC Press, Taylor & Francis Group, 2016).
- Schenkl, S. *et al.* Automatic CT-based finite element model generation for temperature-based death time estimation: Feasibility study and sensitivity analysis. *Int. J. Legal Med.* **131**, 699–712. <https://doi.org/10.1007/s00414-016-1523-0> (2017).
- Weiser, M. *et al.* Uncertainty in temperature-based determination of time of death. *Heat Mass Transf.* **54**, 2815–2826. <https://doi.org/10.1007/s00231-018-2324-4> (2018).
- Hubig, M. *et al.* Fully automatic CT-histogram-based fat estimation in dead bodies. *Int. J. Legal Med.* **132**, 563–577. <https://doi.org/10.1007/s00414-017-1757-5> (2018).
- Fullerton, G. D. Fundamentals of CT Tissue Characterization, In: *Medical Physics of CT and Ultrasound: Tissue Imaging and Characterization* (eds. Fullerton, G. D. & Zagzebski, J. A.) 125–162 (AAPM Medical Physics Monograph No. 6, American Institute of Physics, 1980).
- Buzug, T. M. *Einführung in die Computertomographie: Mathematisch-physikalische Grundlagen der Bildrekonstruktion* (Springer, 2004).
- Welch, B. L. The generalization of “Student’s” problem when several different population variances are involved. *Biometrika* **34**(1–2), 28–35 (1947).
- Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybernet. SMC* **9**(1), 62–66 (1979).
- Mitsiopoulos, N. *et al.* Cadaver validation of skeletal muscle measurement by magnetic resonance imaging and computerized tomography. *J. Appl. Physiol.* **85**, 115–122 (1998).
- Rogalla, P., Meiri, N., Hokschi, B., Boeing, H. & Hamm, B. Low-dose spiral computed tomography for measuring abdominal fat volume and distribution in a clinical setting. *Eur. J. Clin. Nutr.* **52**, 597–602 (1998).
- Yoshizumi, T. *et al.* Abdominal fat: Standardized technique for measurement at CT. *Radiology* **211**(1), 283–286 (1999).
- Glasbey, C. A. & Robinson, C. D. Estimators of tissue proportions from X-ray CT images. *Biometrics* **58**, 928–936 (2002).
- Dobrowolski, A., Romvári, R., Allen, P., Branscheid, W. & Horn, P. Schlachtkörperwertbestimmung beim Schwein – Röntgen-Computertomographie als mögliche Referenzmethode. *Fleischwirtschaft* **84**(3), 109–112 (2004).
- Romvári, R. *et al.* Development of a computed tomographic calibration method for the determination of lean meat content in pig carcasses. *Acta Vet Hung* **54**, 1–10. <https://doi.org/10.1556/AVet.54.2006.1.1> (2006).
- Pednekar, A., Bandekar, A. N., Kakadiaris, I. A., & Naghavi, M. Automatic segmentation of abdominal fat from CT data, in *Proceedings of the Seventh IEEE Workshop on Applications of Computer Vision (WACV/MOTION'05)*, 2005 WACV/MOTIONS '05, vol 1, 308–315 (2005).
- Johansen, J., Egelanddsdal, B., Roe, M., Kvaal, K. & Aastveit, A. H. Calibration models for lamb carcass composition analysis using computerized tomography (CT) imaging. *Chemom. Intell. Lab. Syst.* **87**, 303–311 (2007).
- Kongsro, J., Røe, M., Aastveit, A. H., Kvaal, K. & Egelanddsdal, B. Virtual dissection of lamb carcasses using computer tomography (CT) and its correlation to manual dissection. *J. Food Eng.* **88**, 86–93 (2008).
- McEvoy, F. J., Madsen, M. T., Strathe, A. B. & Svalastoga, E. Hounsfield Unit dynamics of adipose tissue and non-adipose soft tissues in growing pigs. *Res. Vet. Sci.* **84**, 300–304 (2008).
- Ohshima, S. *et al.* Development of an automated 3D segmentation program for volume quantification of body fat distribution using CT. *Jpn. J. Radiol. Technol.* **64**, 1177–1181 (2008).
- McEvoy, F. J., Madsen, M. T., Nielsen, M. B. & Svalastoga, E. Computer tomographic investigation of subcutaneous adipose tissue as an indicator of body composition. *Acta Vet. Scand.* **51**, 28. <https://doi.org/10.1186/1751-0147-51-28> (2009).
- Cecchini, S., Cavazzini, E., Marchesi, F., Sarli, L. & Roncoroni, L. Computed tomography volumetric fat parameters versus body mass index for predicting short-term outcomes of colon surgery. *World J. Surg.* **35**, 415–423. <https://doi.org/10.1007/s00268-010-0888-3> (2011).
- Kanally, C. W. *et al.* A novel method for volumetric MRI response assessment of enhancing brain tumors. *PLoS ONE* **6**(1), e16031. <https://doi.org/10.1371/journal.pone.0016031> (2011).
- Subramaniam, K., Hoffman, E. A., & Tawhai, M. H. Engineering in Medicine and Biology Society (EMBC). In *2012 Annual International Conference of the IEEE San Diego, California USA, 28 August–1 September*, 4072–4089. <https://doi.org/10.1109/EMBC.2012.6345869> (2012).
- Kim, Y. J. *et al.* Body fat assessment method using CT images with separation mask algorithm. *J. Digit. Imaging* **26**, 155–162. <https://doi.org/10.1007/s10278-012-9488-0> (2013).
- Kullberg, J. *et al.* Automated analysis of liver fat, muscle and adipose tissue distribution from CT suitable for large-scale studies. *Sci. Rep.* **7**, 10425. <https://doi.org/10.1038/s41598-017-08925-8> (2017).
- Heinrich, A., Schenkl, S., Buckreus, D., Güttler, F. V. & Teichgräber, U. K. CT-based thermometry with virtual monoenergetic images by dual-energy of fat, muscle and bone using FBP, iterative and deep learning-based reconstruction. *Eur. Radiol.* **66**, 1–8. <https://doi.org/10.1007/s00330-021-08206-z> (2021).
- Schenkl, S. *Quantifizierung des menschlichen Fettgewebes für die temperaturgestützte Todeszeitschätzung mit der Finite-Elemente-Methode*. (Dissertation, University Hospital Jena – Friedrich-Schiller-University Jena, Germany, 2019)
- Jennrich, R. I. & Moore, R. H. Maximum likelihood estimation by means of nonlinear least squares. *J. R. Stat. Soc. B* **59**, 569–587 (1975).
- Papoulis, A. *Probability & Statistics* (Prentice Hall, 1990).

30. Lehmann, E. L. *Theory of Point Estimation* (Wadsworth and Brooks/Cole Advanced Books & Software, 1991).

Acknowledgements

We gratefully acknowledge the aid of Mrs. Antje Kubin as radiographer in the Institute for Diagnostic and Interventional Radiology.

Author contributions

Study design and milestone discussions: All authors; Corpses CTs: Organisation: S.S., F.G., A.H., U.T., G.M.; Perform CTs: S.S., A.H., F.G., J.S.; Data preparation: S.S., F.G., A.H., M.H., H.M.; CFES-Software concept and coding: M.H., S.S., H.M., J.S., A.H., F.G.; Computations and debugging: M.H., S.S., H.M., J.S.; Output interpretation and presentation: M.H., S.S., H.M., A.H.; Paper written and corrected: S.S., M.H.; Paper reading and correction suggestions: All authors.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-24358-4>.

Correspondence and requests for materials should be addressed to M.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022