Enhancing Digital Cephalic Radiography With Mixture Models and Local Gamma Correction

I. Frosio*, *Member, IEEE*, G. Ferrigno, and N. A. Borghese, *Member, IEEE*

***Abstract*—We present a new algorithm, called the soft-tissue filter, that can make both soft and bone tissue clearly visible in digital cephalic radiographies under a wide range of exposures. It uses a mixture model made up of two Gaussian distributions and one inverted lognormal distribution to analyze the image histogram. The image is clustered in three parts: background, soft tissue, and bone using this model. Improvement in the visibility of both structures is achieved through a local transformation based on gamma correction, stretching, and saturation, which is applied using different parameters for bone and soft-tissue pixels. A processing time of 1 s for 5 Mpixel images allows the filter to operate in real time. Although the default value of the filter parameters is adequate for most images, real-time operation allows adjustment to recover under- and overexposed images or to obtain the best quality subjectively. The filter was extensively clinically tested: quantitative and qualitative results are reported here.**

***Index Terms*—Digital radiography, histogram-based clustering, image enhancement, local gamma correction, mixture models, soft-tissue filter (STF).**

I. Introduction

CEPHALIC radiographs are widely used by dentists, surgeons, and maxillofacial radiologists for diagnosis, surgical planning, and implant evaluation [1]. Thanks to modern digital radiographic systems, qualitative evaluation becomes possible in real time, as does quantitative measurement and visualization of anatomical features (e.g., nasal spine, chin tip, ...). Alterations in patient's anatomy and visualization of postoperative aesthetic modifications can be automatically computed [2] and displayed.

To take full advantage of these systems, radiograms are usually treated mathematically so as to obtain optimal grey-level coding, using a variety of techniques, which are generally termed Image Enhancement [3]. The challenge arises from the need to achieve an efficient enhancing solution at interactive rates (processing time $\leq$1 s) for images that are currently on the order of 5 Mpixels.

Manuscript received July 8, 2005; revised October 14, 2005. This work was supported in part by the Gendex Imaging of Cusano Milanino, Milan, Italy. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was E. Krupinsky. *Asterisk indicates corresponding author.*

*I. Frosio is with the Applied Intelligent System Laboratory, Computer Science Department, University of Milan, 20135 Milan, Italy (e-mail: frosio@dsi.unimi.it; website: http://ais-lab.dsi.unimi.it).

G. Ferrigno is with the Bioengineering Department, Politecnico of Milan, 20133 Milan, Italy (e-mail: giancarlo.ferrigno@polimi.it).

N. A. Borghese is with the Applied Intelligent System Laboratory, Computer Science Department, University of Milan, 20135 Milan, Italy (e-mail: borghese@dsi.unimi.it).

Digital Object Identifier 10.1109/TMI.2005.861017

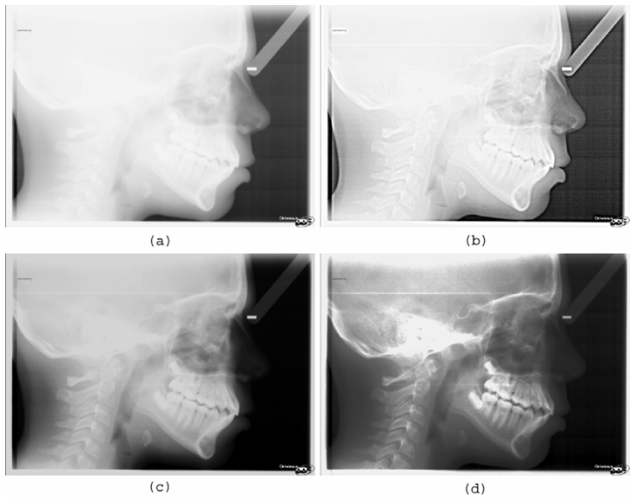

Fig. 1. (a) Typical cephalic radiography, 1871 $\times$ 2605 pixels. (b) Same image treated with UM: the high frequencies are enhanced, but noise is increased, whereas bone is still not clearly visible. (c) Same image after GC, using $\gamma = 0.25$. Although the number of GLs used by the bone pixels (brighter levels) has increased, the range of the dark ones is compressed. (d) Same image after IE. Both GC and IE enhance the bony pixels, but the soft tissue darkens and tends to mix with the background. The same image treated with the proposed filter is shown in Fig. 5(h).

One of the main challenges in cephalometric radiography is to clearly display both soft and bony tissue in the same image [Fig. 1(a)]. Establishing ideal exposure parameters for each patient is very difficult, because of the large difference between the absorption coefficients of the two tissues. In practice, the voltage and the amperage of the X-ray tube are estimated so that the full dynamic range of the X-ray detector is used, taking into account the maximum level of radiation deliverable to the patient. As a result, underexposure of bone and overexposure of soft-tissue often occur, leading to images where the bone and soft-tissue pixels take on similar grey levels (GLs) or the background tends to mix with soft tissue. The substructures inside each tissue then cease to be clearly visible, making their identification difficult if not impossible. The procedures aimed at solving these problems are termed soft-tissue filtering.

A great deal of work has been devoted to making the different structures more visible by increasing the local contrast at the edge of each image element. Unsharp masking (UM) is one of the most widely used techniques, [4]–[8]. It can be implemented to work in real time, but it enhances only the small features of the image and increases the noise. Moreover, it does not allow recover of underexposed images in which the dynamic range of the bone-tissue regions is compressed: the high frequencies in the corresponding regions have too little amplitude to be clearly

visible without adding strong edge artifacts [Fig. 1(b)]. In an overexposed image, UM identifies the bone structures well but cannot recover the soft-tissue boundary, where the transition between soft tissue and background is smooth and poorly defined (large scale); this is critical, for instance, in the chin tip or nose profile.

Scale-space processing [9]–[11] has greater capacity to detect features of different sizes but does not completely solve the problems with UM, especially when large structures are present, as in cephalic images. Different solutions are based on morphological analysis through level sets, morphological operators [12], [13] or anisotropic filtering [14], [15]. Although these approaches guarantee greater homogeneity of the GLs within a given feature, the price paid is computational complexity, which leads to a processing time incompatible with real-time operation. Moreover, they suffer from over- or under-enhancement inside the different regions.

An alternative approach is based on analyzing the histogram to remap the GLs so that the dynamic range both for soft-tissue regions and for bone-tissue regions is maximized. The most widely used technique in clinical practice is global gamma correction (GC) because it can run in real-time. However, no single γ value allows clear visibility of both tissues. The usual setting, which is $\gamma = 0.25$, makes bone structures clearly visible, but soft tissue darkens and tends to mix with the background [Fig. 1(c)]. Gamma values greater than 1.0 can be profitably used to recover overexposed soft tissue but compress the dynamic range in bone regions. Image equalization (IE) produces results very similar to those obtained with GC, $\gamma < 1.0$ [Fig. 1(d)]. The inadequacy of these global approaches is obvious. Therefore, more refined methods that work at a local level have been proposed.

Solutions based on local statistics, such as local histogram equalization [13], [16], [17] or homogeneity analysis [18], [19], reframe the task as a globally constrained nonlinear optimization problem, where the remapping of GLs is constrained at different thresholds while maintaining the same ordering. These solutions have the drawback of being computationally intensive and may suffer from over-enhancement.

We propose here a novel approach to soft-tissue filtering, which is based on identifying soft and bone tissue in the histogram using an appropriate mixture model composed of two Gaussian distributions and one inverted lognormal distribution [20]. A different local transformation, based on GC, linear stretching, and saturation, is then applied to the pixels belonging to the two tissues. The resulting algorithm was widely tested and was consistently able to produce clear visibility of both tissues. Moreover, its processing time of about one second makes this solution fully compatible with the interactive visualization rate required by clinical use.

II. Method

Soft-tissue filtering is obtained by five sequential steps. First, a reliable histogram of the image is built by taking out pixels that belong to borders or to the logotype, as well as saturated pixels. The three components of the histogram (background, soft tissue, and bony tissue) are identified through a mixture model and the optimal threshold between soft and bony tissue is identified. This threshold makes it possible to build a map that contains the GC value for each pixel. Finally, this map is smoothed and applied to the original image.

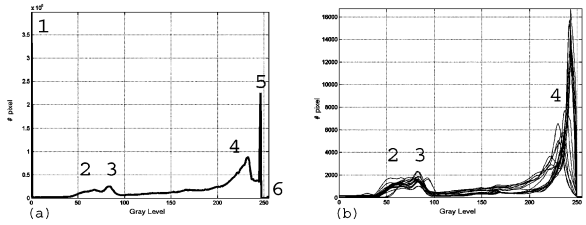

Fig. 2. (a) Typical histogram of a lateral, cephalic radiography shows six marked peaks (see text for a description of the peaks). (b) Histograms of 18 cephalic lateral radiographies after elimination of saturated, edge, and logotype pixels. These were filtered using a moving average filter of size seven.

A. Histogram Description

The histogram of a cephalic radiographic image has a consistent shape (Fig. 2) with six well-defined peaks.

Peak 1 is associated with the pixels that are saturated in the charge-couple device (CCD) sensor, corresponding to GL equal to zero; peaks 2 and 3 represent the image background. The double peak results from automatic exposure control (AEC), which was introduced in the latest generation of radiographic equipment to limit soft-tissue overexposure in the frontal part of the face. Peak 4 is associated with bone structures. It is asymmetrical and shows a steeper slope for the highest GLs. Peak 5 corresponds to pixels at the edge of the CCD sensor, which receive almost no X-rays. Peak 6 is associated with the digital logotype printed on the radiography [corresponding to the maximum GL, equal to $(N_{\mathrm{GL}} - 1)$, where N_{GL} is the number of GLs in the image]. Soft-tissue GLs are spread between peak 2 and peak 4 [Fig. 2(b)]. Under- and overexposed images generate two different histogram populations: as a matter of fact, the bone peak is very high and narrow in underexposed radiographies, whereas it is lower and broader in overexposed ones [Fig. 2(b)].

B. Contour, Saturated Pixels and Logotype Elimination

To obtain a reliable computation for the image histogram, first of all, the pixels on the edge of the image are discarded. A boundary frame as large as 5% of the total number of rows and columns is taken out from the image. This is a safe margin to ensure that all the pixels, that were not fully exposed to radiation are discarded.

At this point, a working histogram of the image, which we denote as $H_{1F}(x)$ for the GL x, is computed, using only the remaining pixels. The peaks associated with saturated pixels (GL equal to zero) and the logotype (GL equal to $\mathrm{N_{GL-1}}$) are now discarded as follows:

$$H_{1F}(0) = H_{1F}(1) \tag{1}$$

$$H_{1F}(N_{\mathrm{GL}} - 1) = H_{1F}(N_{\mathrm{GL}} - 2). \tag{2}$$

The resulting $H_{1F}(\cdot)$ is low-pass filtered using a moving-average filter. It is plotted in Fig. 2(b) for 18 lateral, cephalic radiographies. As can be seen, only peaks 2, 3, and 4 are present in $H_{1F}(\cdot)$. The probability that a certain GL, x, will appear in the

image can be computed by normalizing $H_{1F}(\cdot)$. The resulting normalized histogram will be referred to as $H_{1FN}(\cdot)$.

C. Mixture Model for Histogram Based Clustering

Our purpose is to identify a significant threshold (Th_{Bone}) that allows separating the brighter bone from the darker soft tissue and background pixels. A predefined value cannot be assigned to Th_{Bone} because the levels of the two kinds of tissues, and consequently Th_{Bone}, vary from image to image, depending on the subject's anatomical characteristics.

Therefore, we introduce here a new approach based on mixture models, a powerful statistical technique for estimating probability densities that combines the advantages of both parametric and nonparametric methods [20], [21]. Mixture models can estimate probability densities that have complex shapes, such as multimodal histograms (like the one here), using a restricted number of parameters.

A mixture distribution is defined as a linear combination of M component densities, $p(x|j)$, weighted by the mixing parameters $P(j)$

$$p(x) = \sum_{j=1}^{M} p(x|j) \cdot P(j) \tag{3}$$

with

$$\sum_{j=1}^{M} P(j) = 1 \quad 0 \le P(j) \le 1 \tag{4}$$

$$\int p(x|j)dx = 1, \ \forall j. \tag{5}$$

In practice, the probability density $p(x)$ is generated as follows: first a component j is chosen with probability $P(j)$, then a data point is generated from the component density $p(x|j)$. Posterior probabilities can be expressed using Bayes' theorem as

$$P(j|x) = \frac{p(x|j)P(j)}{p(x)} \tag{6}$$

with

$$\sum_{j=1}^{M} P(j|x) = 1 \tag{7}$$

where $P(j|x)$ is the probability that a particular component j has generated x.

Analysis of the histogram shown in Fig. 2 has inspired us to model H_{1FN} as a mixture of two Gaussian distributions and one inverted lognormal distribution, where each component in the mixture takes into account the spread of the GLs associated with background, soft-tissue, and bone, respectively. The characteristic shape of the inverted lognormal distribution is used to properly describe the asymmetric shape of the bone peak. This particular distribution has different formulations: the probability density of the one used here is given by

$$p(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma \cdot (N_{\mathrm{GL}} - x)} \cdot \exp\left\{ -\frac{[\ln(N_{\mathrm{GL}} - x) - \mu]^2}{2\sigma^2} \right\} \tag{8}$$

and it is defined only for $x < N_{\mathrm{GL}}$. Its mean and variance are, respectively

$$M = N_{\mathrm{GL}} - 1 - \exp\left(\mu + \frac{\sigma^2}{2}\right) \tag{9}$$

$$S^2 = \exp(\sigma^2 + 2\mu) \cdot \left(\exp(\sigma^2) - 1\right). \tag{10}$$

The other components of the mixture model make up a Gaussian distribution, described by

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{11}$$

where μ and σ^2 are the mean and the variance of the distribution.

The mixture model is, thus, completely defined, once the parameters of each distribution (μ_j, σ_j) and the three mixing parameters $P(j)$ have been computed.

A common method for determining these parameters is to maximize the likelihood function of the parameters for the given dataset. More easily, we can minimize the negative log-likelihood, given by

$$E = -\ln L = -\sum_{n=1}^{N} \ln p(x^n) = -\sum_{n=1}^{N} \ln \{p(x^n|j)P(j)\} \tag{12}$$

where N is the size of the dataset.

A closed-form solution for computing the parameters by minimizing E in (12) is not available, so iterative algorithms have to be adopted. A common solution is represented by the expectation-maximization (EM) algorithm [20], [21], which produces the equations for updating the parameters at each iteration step. Deriving EM solutions for standard distributions like Gaussians and Poisson distributions can be found in [20], [21]. The particular model used here made it necessary to derive the updating equations reported in Appendix A.

A huge computational task would, thus, be created, because each pixel should be examined using (A12)–(A14), (A17), and (A18). However, the possible GLs for each pixel are discrete and finite (N_{GL} values). Moreover, all the pixels with the same grey value have already been counted in histogram H_{1F}. The updating equations can then be simplified as

$$P(j)^{\mathrm{new}} = \frac{1}{N} \sum_{g=0}^{N_{\mathrm{GL}}-1} P^{\mathrm{old}}(j|g) \cdot H_{1F}(g) \tag{13}$$

for each j,

$$\mu_j^{\mathrm{new}} = \frac{\sum_{g=0}^{N_{\mathrm{GL}}-1} P^{\mathrm{old}}(j|g) \cdot g \cdot H_{1F}(g)}{\sum_{g=0}^{N_{\mathrm{GL}}-1} P^{\mathrm{old}}(j|g) \cdot H_{1F}(g)} \tag{14}$$

$$\left(\sigma_j^{\mathrm{new}}\right)^2 = \frac{\sum_{g=0}^{N_{\mathrm{GL}}-1} P^{\mathrm{old}}(j|g) \cdot \left(g - \mu_j^{\mathrm{new}}\right)^2 \cdot H_{1F}(g)}{\sum_{g=0}^{N_{\mathrm{GL}}-1} P^{\mathrm{old}}(j|g) \cdot H_{1F}(g)} \tag{15}$$

for the two Gaussian distributions, and

$$\mu_j^{\text{new}} = \frac{\sum_{g=0}^{N_{\text{GL}}-1} P^{\text{old}}(j|g)\cdot \ln(N_{\text{GL}}-g)\cdot H_{1F}(g)}{\sum_{g=0}^{N_{\text{GL}}-1} P^{\text{old}}(j|g)\cdot H_{1F}(g)} \quad (16)$$

$$\left(\sigma_j^{\text{new}}\right)^2 = \frac{\sum_{g=0}^{N_{\text{GL}}-1} P^{\text{old}}(j|g)\cdot \left[\ln(N_{\text{GL}}-g)-\mu_j^{\text{new}}\right]^2\cdot H_{1F}(g)}{\sum_{g=0}^{N_{\text{GL}}-1} P^{\text{old}}(j|g)\cdot H_{1F}(g)} \quad (17)$$

for the inverted lognormal distribution.

We explicitly note that the term $\sum_{g=0}^{N_{\text{GL}}-1} P^{\text{old}}(j|g)\cdot H_{1F}(g)$ is common to the equations updating the mixing parameters (13) and the component parameters [(14), (15) or (16), (17)]. Therefore, this term needs to be computed only once for each updating step.

To obtain a reliable estimate and maximize convergence speed, the parameters are initialized to their mean value, as computed for a set of beta-test images.

After the mixture-model parameters have converged, the first Gaussian component corresponds to the background; the second is associated with soft tissue; and the third component, the inverted lognormal distribution, describes bone-tissue. The convergence of the model can be observed in Fig. 3(a)–(d). Fig. 3(e) shows how E is minimized when the mixture model is being trained. The threshold that separates soft tissue from bone structures, Th_{Bone}, can be set so that the following function is minimized:

$$\int_0^{Th_{\text{Bone}}} p(x|3)P(3)dx + \int_{Th_{\text{Bone}}}^{N_{\text{GL}}} p(x|2)P(2)dx \quad (18)$$

that is, the probability of assigning x to the wrong component j is minimized, for $j=2$ or $j=3$.

Finally, the greatest significant GL of the image, $G_{\max}$, can then be identified as

$$G_{\max} = \min(N_{\text{GL}}-1, M+S) \quad (19)$$

where M and S are the mean and the standard deviations, respectively, of the inverted lognormal distribution of the mixture model. Th_{Bone} and $G_{\max}$ change their values as the mixture model is trained, which can be observed in Fig. 3(f).

D. Gamma Map and Local Gamma Correction

At this stage, we could apply pixel-to-pixel GC as

$$I'(i,j) = (N_{\text{GL}}-1)\left[\frac{I(i,j)}{N_{\text{GL}}-1}\right]^{\frac{1}{\gamma(i,j)}} \quad (20)$$

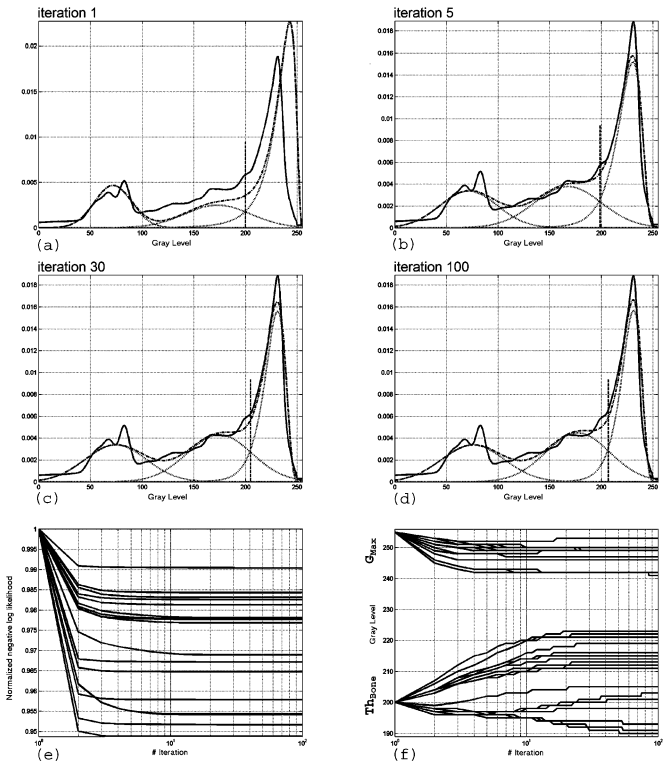

Fig. 3. Normalized histogram of the original image, H_{1FN}, is plotted with "___," the probability density of each GL, computed by the mixture model, with "-.-.-.," and the probability density of the three components, with ". . .." They are plotted at iteration step 1 (a), 5 (b), 30 (c), and 100 (d) of the EM algorithm. Computed Th_{Bone} is shown in each diagram as a vertical dashed line. The negative log-likelihood, E, normalized to its initial value, against the number of iteration steps for 18 beta-test radiographies, is plotted in panel (e). Value of the threshold, Th_{Bone}, and the maximum significant GL, $G_{\max}$, are plotted in (f).

where $I(i,j)$ is the GL of the pixel (i,j) in the original image and $I'(i,j)$ is its value in the image transformed by $\gamma(i,j)$. Specifically, each pixel (i,j) in which $I(i,j) \le Th_{\text{Bone}}$ will be modified by using $\gamma(i,j) = \gamma_{\text{Soft\_tissue}}$, whereas $\gamma(i,j) = \gamma_{\text{Bone}}$ will be used for pixels (i,j) in which $I(i,j) > Th_{\text{Bone}}$. Although this procedure is fast, it does not take into account the fact that the sharpness of the transition between bone and soft tissue is usually grater than one pixel. Therefore, γ values have to be smoothed in the spatial domain to avoid strong artifacts, as shown in Fig. 4(e).

Therefore, we first create a binary gamma map, $\Gamma_b(.)$, which contains either the value $\gamma_{\text{Soft\_tissue}}$ or γ_{Bone} [Fig. 4(b)]. $\Gamma_b(.)$ has to be spatially filtered to obtain the final gamma map, $\Gamma_f(.)$ [Fig. 4(d)], which will be applied to the image.

First $\Gamma_b(.)$ is down-sampled into $\Gamma_d(.)$. For this purpose, $\Gamma_b(.)$ is subdivided into square blocks, $B_{l,m}$, whose size is $TP \times TP$. $\Gamma_d(l,m)$ contains the mean gamma value inside block, $B_{l,m}$. $\Gamma_d(.)$ is then spatially filtered using a 3 $\times$ 3 moving-average window [Fig. 4(c)]. This procedure is equivalent to undersampling $\Gamma_b(.)$ with partially overlapping windows. Lastly, $\Gamma_f(.)$ is obtained by upsampling $\Gamma_d(.)$ through a bilinear interpolation scheme [Fig. 4(d)].

To take advantage of the full dynamics of the GLs, linear stretching with saturation is applied to histogram H_{1F}, before local GC. Combining linear stretching with saturation and with

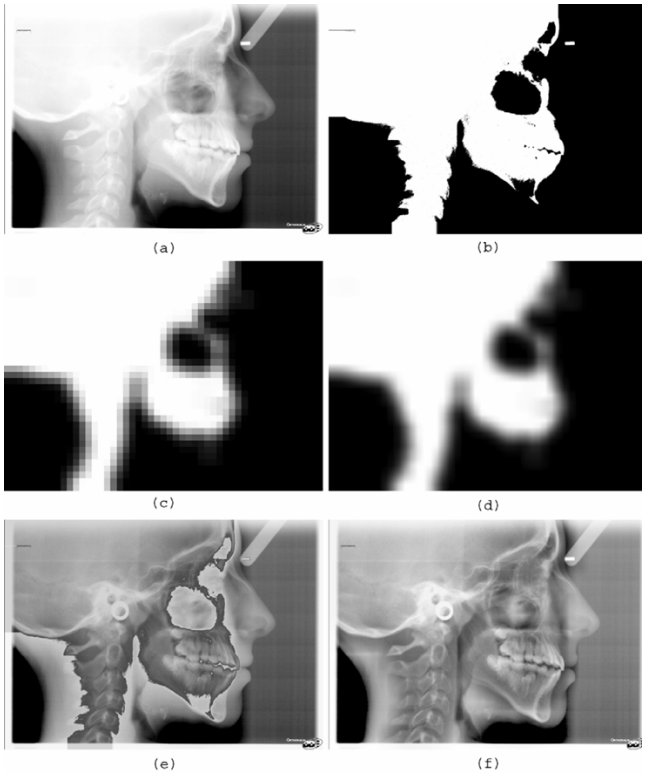

Fig. 4. (a) Cephalic radiography. (b) Gamma binary field $\Gamma_b(.)$ extracted from the same image. (c) Gamma field undersampled and filtered, $\Gamma_d(.)$. (d) The final gamma map used for image correction, $\Gamma_f(.)$. (e) The image filtered with binary gamma map, $\Gamma_b(.)$; artifacts are evident. (f) The result obtained by applying $\Gamma_f(.)$, using $\gamma_{\mathrm{Bone}} = 0.25$, $\gamma_{\mathrm{SoftTissue}} = 2.2$, and $TP = N_{\mathrm{ROW}}/24$.

local GC (20) yields this final correction formula for each pixel (i,j)

$$I'(i,j) = \begin{cases} (N_{\mathrm{GL}}-1)\cdot\left[\frac{I(i,j)}{G_{\max}}\right]^{\frac{1}{\Gamma_f(i,j)}}, & I(x,y) < G_{\max} \\ (N_{\mathrm{GL}}-1), & \text{elsewhere} \end{cases}. \tag{21}$$

Maximum speed-up of the algorithm is achieved by implementing (21) through a lookup table (LUT). The map $\Gamma_f(i,j)$ is discretized into N_V values, $\Gamma_0,\ldots,\Gamma_{N_V-1}$. For each GL, I_p, $0 \le p \le N_{\mathrm{GL}}-1$, and for each gamma value, Γ_k, $0 \le k \le N_V - 1$, the corrected level, I'_p, is computed through (21) and stored in the LUT, whose size is, thus, $N_{\mathrm{GL}} \times N_V$. Each pixel $I(i,j)$ is then corrected by accessing the LUT table in position $[XI(i,j), \Gamma(i,j)]$

$$I'(i,j) = LUT\left[I(i,j), \Gamma(i,j)\right]. \tag{22}$$

E. Algorithm Implementation and Evaluation

Qualitative and quantitative results will be reported for a beta-test set made up of 18 lateral 1871 × 2606 pixels cephalometric images, acquired with the Gendex Orthoralix 9200 DDE. The proposed algorithm was compared with IE, GC, and UM, techniques widely used in clinical practice. After beta testing, the filter was distributed and has so far been used by more than 300 dentists.

The soft-tissue filter (STF) algorithm was implemented in C++ on an Athlon 2400+ at 1.79 GHz with 256 Mb RAM. The same machine was used to measure processing time for the beta-test images.

In addition to qualitative evaluation, quantitative assessment of STF was carried out through two indexes: the Shannon entropy index and a local contrast enhancement index.

Shannon entropy [22] is a measure of the information contained in the image and it defined as

$$H = -\sum_{x=0}^{N_{\mathrm{GL}}-1} p(x)\log_2\left(p(x)\right) \tag{23}$$

where $p(x)$ is the probability of GL x occurring in the image. In order to compare the processing effect on a population of images with widely divergent H values, normalized entropy, H_N, was adopted. This is defined as the ratio of the entropy of the treated image to that of the original image.

Local contrast enhancement was computed as follows. For each of the 18 beta-test images, an 80 × 80 pixel square, centered on the left-most molar in the image, was manually selected. The local contrast inside this square was evaluated both for the original images and for each original treated with IE, GC, UM, and the proposed filter. According to [23], local contrast C is measured as

$$C = \frac{o-b}{b} \tag{24}$$

where o and b are defined as the 75th and 25th percentiles, respectively, of the GLs inside the square. The contrast values were normalized to the ones computed for the original image, so as to obtain a measurement of the local contrast enhancement that was provided by each method for each image. The normalized contrast values will be referred to as C_N.

III. Results

For each radiography of our beta-test dataset, the filter improved the visibility of both bony and soft tissue anatomical structures on the same image; this was achieved under a wide range of exposures, including underexposed and overexposed radiographies. Typical results are shown in Fig. 5(a)–(d).

The result of applying STF to overexposed images is shown in Fig. 5(e), (f). Here, the contrast between bone and soft tissue is quite high, but soft tissue tends to merge with the background. After STF treatment, the bone structures become more visible, so chin and soft tissue can be clearly distinguished from the background [Fig. 5(f)]. The use of greater-than-default filter size (TP) makes the transition between $\gamma_{\mathrm{Soft\_tissue}}$ and γ_{Bone} in Γ_f very smooth.

Similar results were obtained for underexposed radiographies [Fig. 5(g), (h)]. Here, soft tissue is visible, but the grey-level range of the bone regions is very compressed. Moreover, the contrast between soft tissue and bone is very low. Treatment with STF, using low γ_{Bone} $(\gamma_{\mathrm{Bone}} = 0.15)$ increases bone-structure visibility, while leaving soft tissue almost untouched [Fig. 5(h)].

The run time for the algorithm was one second on each image. Specifically, for images of 1871 by 2606 pixels, processing time,

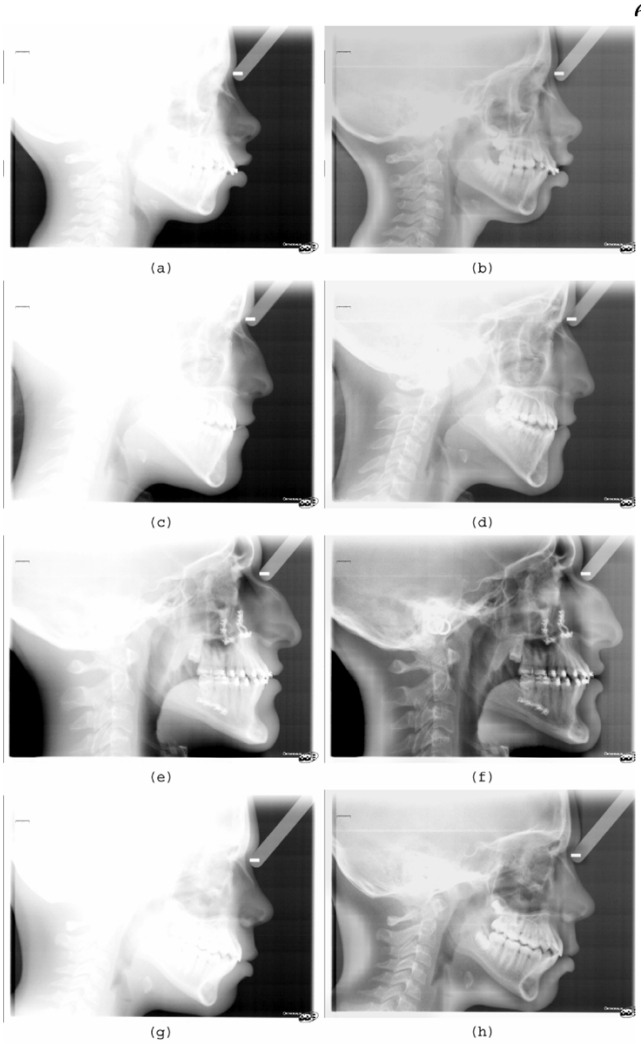

Fig. 5. Lateral cephalic radiographies: (a) and (c) originals and (b) and (d) those treated with STF. Default settings were used: $\gamma_{\mathrm{Bone}} = 0.25$, $\gamma_{\mathrm{Soft\ Tissue}} = 1.5$, $T_P = N_{\mathrm{ROW}}/36$, where N_{ROW} is the number of rows of the image. An overexposed, lateral cephalic radiography, (e) original and (f) one treated with STF. Settings were: $\gamma_{\mathrm{Bone}} = 0.2$, $\gamma_{\mathrm{Soft\ Tissue}} = 2$, $TP = N_{\mathrm{ROW}}/24$. An underexposed, lateral cephalic radiography, (g) original and (h) treated with STF. Settings were: $\gamma_{\mathrm{Bone}} = 0.15$, $\gamma_{\mathrm{Soft\ Tissue}} = 1$, $TP = N_{\mathrm{ROW}}/36$.

T_p, was 1.08 ± 0.01s (mean $\pm$ 2 standard deviations): only 6% of T_p is devoted to computing the histogram and analyzing it using the mixture model. Constructing Γ_b, smoothing it to obtain Γ_f, and grey-level correction using the LUT take up 17%, 42%, and 17% of T_p, respectively. The remaining 18% of T_p is used for allocating memory. These times make it possible to work at interactive rates and adjust the parameters (γ_{Bone}, $\gamma_{\mathrm{Soft\_Tissue}}$ and TP) to obtain optimal results subjectively.

H_N values (mean $\pm$ 2 standard deviation) are reported in Fig. 6(a), for 18 radiographies treated with IE, GC, UM, and STF. As can be seen, IE (0.81 ± 0.02) and GC (0.87 ± 0.04) produce a decrease in the image's information content, whereas UM increases it by a small amount (1.04 ± 0.04). On the other hand, STF leaves unaltered (0.98 ± 0.06) the information contained in the image.

The C_N values (mean $\pm$ 1 standard deviation) are reported in Fig. 6(b) for the images treated with IE, GC, UM, and STF.

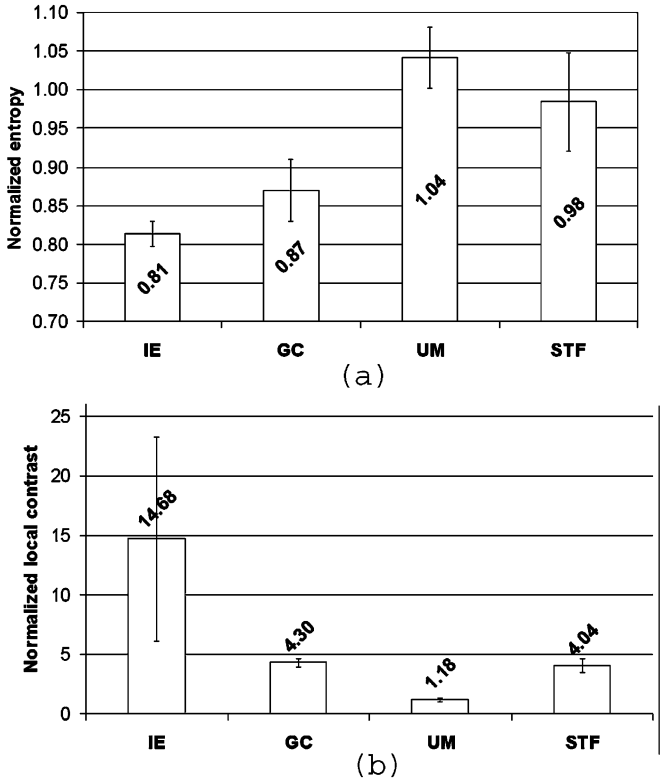

Fig. 6. (a) Mean normalized entropy, H_N, ± 2 standard deviations, is reported for the beta test set of 18 images, treated with IE, GC ($\gamma = 0.25$), UM ($\mathrm{mask\ dimension} = 25 \times 25, \mathrm{gain} = 2$), and STF ($\gamma_{\mathrm{Soft\_tissue}} = 1.5$, $\gamma_{\mathrm{Bone}} = 0.25, TP = N_{\mathrm{ROW}}/36$). (b) Mean normalized local contrast, C_N, ± 1 standard deviation, is reported for the same image dataset.

The greatest enhancement of local contrast (14.68 ± 8.62) is obtained through IE; GC and STF lead to a similar increase in local contrast: 4.30 ± 0.36 and 4.04 ± 0.60, respectively. UM gives rise to the smallest enhancement: 1.18 ± 0.13.

IV. Discussion

A. Histogram Analysis Through the Mixture Model

The core of STF is an innovative modeling of the histogram that uses an appropriate mixture model, which has allowed a reliable clustering of the cephalic images, in practically negligible time (less than 60 ms).

To achieve this result the components of the mixture have to be carefully chosen [21]. By visual inspection, four different grey-level distributions seem to be present in cephalic images. Two of these are associated with the background [peaks 2 and 3 in Fig. 2(a)], where AEC is turned on or off. Aside from the fact that we were not interested in distinguishing between these two distributions, they are not statistically significantly different. Thus, the use of four components produces over-fitting and makes the solution unstable: the fourth component does not model any peak of the histogram and varies greatly in position and in amplitude from image to image. Therefore, the two background distributions are modeled with a single Gaussian distribution.

Soft tissue also can be represented with a Gaussian distribution of GLs. On the other hand, the histogram associated with bone tissue is more complex. As a matter of fact, this has a

markedly asymmetrical shape and exhibits peaks of varying position and amplitude, depending on the exposure parameters: underexposure increases peak amplitude and the peak position, while decreasing peak width.

The inverted lognormal distribution, adopted to describe the third peak of the histogram, has been shown to work properly, both on underexposed radiographies, which have a narrow bone peak, and on overexposed images, with a more spread out histogram. Overall, the three components represent background, soft tissue and bony tissue well, as shown in Fig. 3(d). Among the possible analytical shapes of the inverted lognormal distribution, the one adopted here was chosen because it enabled us to derive closed analytical equations to update the parameters. This form is suitable for use with the EM algorithm, which produces a fast and stable convergence [20], [21].

In the beta-test images, no change in the GL associated with Th_{Bone} and $G_{\max}$ was observed after 100 iterations [Fig. 3(f)]. In as few as 50 iterations, the negative log-likelihood stops decreasing significantly, as shown in Fig. 3(e). The same is true for Th_{Bone} and $G_{\max}$, which may change one GL at most, for some images after 50 iterations. Initialization is not critical: positioning the three components equally spaced in the GL domain, and setting the variance of the three components to $N_{\mathrm{GL}}/10$, increases computing time by less than 6 ms (the number of iterations to achieve the same figures increases by less than 10%).

B. Filter Description

The transformation (21) was optimized. It is applied to each pixel of the radiography and is composed of linear stretching, saturation, and local GC. Due to the saturation component, all the pixels whose GLs are higher than $G_{\max}$, are clipped to $N_{\mathrm{GL}} - 1$ in the filtered image. As $G_{\max}$ is properly estimated by the mixture, only almost empty GLs are eliminated. On the other hand, these levels are recovered by stretching; thus leading to an increase in image contrast.

Three parameters are associated with STF: $\gamma_{\mathrm{Soft\_Tissue}}$, γ_{Bone} and TP. Their value is set by default at: $\gamma_{\mathrm{Bone}} = 0.25$, $\gamma_{\mathrm{Soft\_Tissue}} = 1.5$, $TP = N_{\mathrm{ROW}}/36$, which allows good results to be obtained over a wide variety of images, as demonstrated in Fig. 5(a)–(d). However, thanks to the computational efficiency of the proposed method, the user can modify these values interactively to obtain the best image quality subjectively. Moreover, this also allows recovery of highly under- or overexposed images; underexposure can be corrected by using a small value for γ_{Bone} [for instance, $\gamma_{\mathrm{Bone}} = 0.15$ as in Fig. 5(g), (h)], while a high value of $\gamma_{\mathrm{Soft\_Tissue}}$, allows recovery of overexposed soft tissue [for instance $\gamma_{\mathrm{Soft\_Tissue}} = 2$ in Fig. 5(e), (f)].

To take into account the fact that the transition between bone and soft tissue is usually not as sharp as one pixel, γ is smoothed in the spatial domain to avoid the strong artifacts shown in Fig. 4(e). The smoothing effect can be modulated by the user: increasing TP enlarges the transition zones. In the proposed algorithm, we then adopted a four-step procedure, which leads to a bilinear interpolation scheme to generate Γ_f. We avoided using spline-based upsampling, which can introduce oscillations in Γ_f. The binary map Γ_b could also be smoothed by a simple moving-average filter, which can be implemented efficiently in the spatial domain so as to work in real time. However, this solution generates artifacts in Γ_f, due to the square shape of the moving window. More refined techniques, aimed at selecting the filter scale locally, on the basis of scale-space analysis, require more computing time, and are not justified here.

The filter proposed here is equivalent to a local nonlinear stretching of the grey-level dynamic range: both soft-tissue and bone dynamic ranges are enlarged, leading to increased visibility of anatomical structures. In the filtered images, bone and soft tissue may share the same GLs, as can be seen in Fig. 5. Although this may be critical for some applications, such as analyzing bone density or investigating tumors, it holds little importance when the clinician has to identify anatomical features precisely, locate alterations in the patient's anatomy or visualize postoperative aesthetic modifications.

C. Quantitative Evaluation

The quantitative evaluation of image processing results is still subject to debate. A large number of quantitative descriptors have been introduced to calculate different image characteristics, but there is not yet complete agreement between these indexes and the evaluation obtained by visual inspection [14]. In our case, where the aim of STF is to make both bone and soft tissue filter clearly visible in a cephalic radiographic image without introducing any artifact, Shannon entropy seems an adequate index [14]. It quantifies the information contained in an image, taking in account only the distribution of the GLs. We also measured local contrast enhancement for some of the techniques most widely used clinically. As a matter of fact, contrast is a critical parameter for the visibility of an object (i.e., any anatomical structure): increasing the contrast between an object and the background improves the object's visibility, as stated by Weber's law [23].

As shown in Fig. 6(a), both IE and GC decrease image information. This is borne out by visual inspection of the processed image (Fig. 1): soft tissue GLs are compressed and become less visible after transformation; the contrast in bone structures is enhanced, as demonstrated in Fig. 6(b), but only a few GLs are used by the bony pixels. Although the increase in local contrast can be high, the loss of GLs produces artifacts like soft-tissue mixing with the background, which make this filter useless in practice. UM, on the contrary, being fundamentally a derivative technique, introduces noise and edge artifacts in the treated radiograms, leading to a small increase in information. Moreover, local contrast enhancement provided by UM is very small, because the effect of this filter is significant only in regions close to the edges. STF, on the other hand, leaves information content unaltered [Fig. 6(a)] and, at the same time, significantly increases local contrast [Fig. 6(b)]. These two characteristics improve the visibility of the anatomical structures of the image, as is apparent from visual inspection (Fig. 5).

The feedback obtained from dentists and maxillofacial surgeons during clinical trials was very positive. All of them observed a great increase in the readability of radiographies filtered using standard parameters. Moreover, they appreciated being able to modify the filter parameters interactively so as to obtain the best visualization results.

V. Conclusion

The filtering algorithm reported here has been widely tested in clinical routines and has proven a powerful tool for visualizing both soft tissue and bone in the same image clearly. Moreover it can be integrated with the latest tools for automatic cephalometric orthodontics [2]. The speed of operation and the intuitive modification of free parameters make it a handy tool for its users. The approach described here can be adapted to all other types of medical images that are characterized by a well-defined multi modal histogram, for which the different tissues have to be displayed clearly in the same image.

Appendix A

For a mixture model made up of M components, $p_{\boldsymbol{\theta}_j}(x|j)$, trained using N data points, the negative logarithm of the likelihood function is given by

$$E = -\ln L = -\ln \prod_{n=1}^{N} p_{\boldsymbol{\theta}}(x^n) = -\sum_{n=1}^{N} \ln p_{\boldsymbol{\theta}}(x^n) \tag{A1}$$

where $\boldsymbol{\theta_j}$ represents the vector of parameters of the jth mixture, $\boldsymbol{\theta}$ represents the ensemble of all the parameters: $\boldsymbol{\theta} = \{\boldsymbol{\theta_j}\}$ and x^n is the nth data point. For simplicity of notation, we will omit $\boldsymbol{\theta}$ in the following.

Our task is to estimate the value of $\boldsymbol{\theta}$ by minimizing E, (A1). The following iterative procedure (EM algorithm) can be employed. First, we can compare the negative logarithm of the likelihood function, before and after updating $\boldsymbol{\theta}$, and obtain

$$\begin{aligned} E^{\text{new}} - E^{\text{old}} &= -\sum_{n=1}^{N} \ln p^{\text{new}}(x^n) - \left(-\sum_{n=1}^{N} \ln p^{\text{old}}(x^n) \right) \\ &= -\sum_{n=1}^{N} \ln \left\{ \frac{p^{\text{new}}(x^n)}{p^{\text{old}}(x^n)} \right\} \end{aligned} \tag{A2}$$

where "old" and "new" indicate the values before and after each iteration step.

Bearing in mind that $p(x) = \sum_{j=1}^{M} P(j) \cdot p(x|j)$ (3), (A2) can be rewritten as

$$\begin{aligned} &E^{\text{new}} - E^{\text{old}} \\ &= -\sum_{n=1}^{N} \ln \left\{ \frac{1}{p^{\text{old}}(x^n)} \cdot \sum_{j=1}^{M} [P^{\text{new}}(j) p^{\text{new}}(x^n|j)] \right\} \\ &= -\sum_{n=1}^{N} \ln \left\{ \sum_{j=1}^{M} \left[\frac{P^{\text{new}}(j) p^{\text{new}}(x^n|j)}{p^{\text{old}}(x^n)} \cdot \frac{P^{\text{old}}(j|x^n)}{P^{\text{old}}(j|x^n)} \right] \right\} \end{aligned} \tag{A3}$$

where the last factor inside squared brackets is simply one.

Given a set of M numbers λ_j, such that $\sum_{j=1}^{M} \lambda_j^2 = 1$, Jensen's inequality [20] says that

$$\ln \left(\sum_{j=1}^{M} \lambda_j^2 K_j \right) \geq \sum_{j=1}^{M} \lambda_j^2 \ln(K_j). \tag{A4}$$

Since we have

$$\sum_{j=1}^{M} P^{\text{old}}(j|x^n) = 1, \ \forall n \tag{A5}$$

we can substitute $P^{\text{old}}(j|x^n)$ to λ_j^2 and $(\sum_{j=1}^{M} P^{\text{new}}(j) p^{\text{new}}(x^n|j) / p^{\text{old}}(x^n)) \cdot (1/P^{\text{old}}(j|x^n))$ to K_j in (A4), and we can write

$$\begin{aligned} &\ln \left\{ \sum_{j=1}^{M} \left[P^{\text{old}}(j|x^n) \cdot \frac{P^{\text{new}}(j) p^{\text{new}}(x^n|j)}{p^{\text{old}}(x^n) \cdot P^{\text{old}}(j|x^n)} \right] \right\} \\ &\qquad \geq \sum_{j=1}^{M} P^{\text{old}}(j|x^n) \ln \left\{ \frac{P^{\text{new}}(j) p^{\text{new}}(x^n|j)}{p^{\text{old}}(x^n) \cdot P^{\text{old}}(j|x^n)} \right\}. \end{aligned} \tag{A6}$$

Substituting then (A6) in (A3)

$$\begin{aligned} &E^{\text{new}} - E^{\text{old}} \leq \\ &\quad -\sum_{n=1}^{N} \sum_{j=1}^{M} P^{\text{old}}(j|x^n) \ln \left\{ \frac{P^{\text{new}}(j) p^{\text{new}}(x^n|j)}{p^{\text{old}}(x^n) \cdot P^{\text{old}}(j|x^n)} \right\}. \end{aligned} \tag{A7}$$

Defining Q as the right-hand side in (A7), we see that $E^{\text{new}} \leq E^{\text{old}} + Q$. Minimizing Q, E_{new} is also minimized. Because E_{new} is a function of the new mixture parameters, $\boldsymbol{\theta}^{\mathbf{new}}$, the minimization of Q is carried out with respect to the "new" values of the parameters.

Moreover, because $\boldsymbol{\theta}^{\mathbf{old}}$ is given inside an iteration step, $p^{\text{old}}(x^n) \cdot P^{\text{old}}(j|x^n)$ is fixed and can be taken out of Q. Thus, the following expression can be minimized:

$$\widetilde{Q} = -\sum_{n=1}^{N} \sum_{j=1}^{M} P^{\text{old}}(j|x^n) \ln \{ P^{\text{new}}(j) p^{\text{new}}(x^n|j) \}. \tag{A8}$$

In this minimization, the following constraint must be satisfied:

$$\sum_{j=1}^{M} P^{\text{new}}(j|x^n) = 1, \ \forall n \tag{A9}$$

and the *Lagrangian multiplier* method can be adopted. To find $\boldsymbol{\theta}^{\mathbf{new}}$, we minimize the following constrained expression:

$$f = \widetilde{Q} + \psi \left(\sum_{j=1}^{M} P^{\text{new}}(j) - 1 \right) \tag{A10}$$

where ψ is a the *Lagrangian multiplier*.

The updating equations of EM can then be computed by

$$\begin{cases} \frac{\partial f}{\partial \mu_j^{\text{new}}} = 0 \\ \frac{\partial f}{\partial \sigma_j^{\text{new}}} = 0 \\ \frac{\partial f}{\partial P^{\text{new}}(j)} = 0 \end{cases}. \tag{A11}$$

The system (A11) has nine equations and ten unknowns (the nine mixture parameters plus the Lagrangian multiplier). In order make it computable, (A5) has to be added.

Solving (A11) and (A5) for $P^{\text{new}}(j)$ results in, [20]

$$P^{\text{new}}(j) = \frac{1}{N}\sum_{n=1}^{N} P^{\text{old}}(j|x). \tag{A12}$$

The mean and variance of the Gaussian components can be updated as, [20]

$$\mu_j^{\text{new}} = \frac{\sum_{n=1}^{N} P^{\text{old}}(j|x^n)x^n}{\sum_{n=1}^{N} P^{\text{old}}(j|x^n)} \tag{A13}$$

$$\left(\sigma_j^{\text{new}}\right)^2 = \frac{\sum_{n=1}^{N} P^{\text{old}}(j|x^n)(x^n-\mu_j)^2}{\sum_{n=1}^{N} P^{\text{old}}(j|x^n)}. \tag{A14}$$

For the inverted lognormal distribution, the updating equations for parameters μ_j^{new} and σ_j^{new} can be derived as follows. Solving (A11) for the parameters of the inverted lognormal distributions, the following equations are obtained:

$$\frac{\partial f}{\partial \mu_j^{\text{new}}} = -\sum_{n=1}^{N} P^{\text{old}}(j|x^n)\cdot\left\{\frac{\ln(N_{\text{GL}}-x^n)-\mu_j^{\text{new}}}{\left(\sigma_j^{\text{new}}\right)^2}\right\} = 0 \tag{A15}$$

$$\frac{\partial f}{\partial \sigma_j^{\text{new}}} = -\sum_{n=1}^{N} P^{\text{old}}(j|x^n)\cdot\left\{\frac{-\left(\sigma_j^{\text{new}}\right)^2+\left[\ln(N_{\text{GL}}-x^n)-\mu_j^{\text{new}}\right]^2}{\left(\sigma_j^{\text{new}}\right)^2}\right\} = 0. \tag{A16}$$

The parameters μ_j^{new} and σ_j^{new} can be obtained by multiplying the updating (A15) and (A16) by $\left(\sigma_j^{\text{new}}\right)^2$ and $\left(\sigma_j^{\text{new}}\right)^3$, respectively

$$\mu_j^{\text{new}} = \frac{\sum_{n=1}^{N} P^{\text{old}}(j|x^n)\ln(N_{\text{GL}}-x^n)}{\sum_{n=1}^{N} P^{\text{old}}(j|x^n)} \tag{A17}$$

$$\left(\sigma_j^{\text{new}}\right)^2 = \frac{\sum_{n=1}^{N} P^{\text{old}}(j|x^n)\cdot\left[\ln(N_{\text{GL}}-x^n)-\mu_j^{\text{new}}\right]^2}{\sum_{n=1}^{N} P^{\text{old}}(j|x^n)}. \tag{A18}$$

ACKNOWLEDGMENT

The authors wish to thank P. Grew for reviewing the manuscript.

REFERENCES

[1] R. E. Moyers, *Handbook of Orthodontics*. Chicago, IL: Year Book, 1988.
[2] Y. Chen, K. Cheng, and J. Lin, "Feature subimage extraction for cephalogram landmarking," in *Proc. IEEE EMBS*, 1998, pp. 1414–1417.
[3] B. Chanda and D. D. Majumder, *Digital Image Processing and Analisys*. New Delhi, India: Prentice-Hall of India Ptv. Ltd., 2000.
[4] T.-H. Lin and T. Kao, "Adaptive local contrast enhancement method for medical images diplayed on a video monitor," *Med. Eng. Phys.*, vol. 22, pp. 79–87, Mar. 2000.
[5] D.-C. Chang and W.-R. Wu, "Image contrast enhancement based on a histogram transformation of local standard deviation," *IEEE Trans. Med. Imag.*, vol. 17, no. 8, pp. 518–531, Aug. 1998.
[6] T-.L. Ji, M. K. Sundarseshan, and H. Roehrig, "Adaptive image contrast enhancement based on human visual properties," *IEEE Trans. Med. Imag.*, vol. 13, no. 4, pp. 573–586, Dec. 1994.
[7] A. Polesel, G. Ramponi, and V. J. Mathews, "Image enhancement via adaptive unsharp masking," *IEEE Trans. Image Process.*, vol. 9, no. 3, pp. 505–510, Mar. 2000.
[8] P. G. Tahoces, J. Rocca, M. Souto, C. Gonzalez, L. Gòmez, and J. J. Vidal, "Enhancement of chest and breast radiographs by automatic spatial filtering," *IEEE Trans. Med. Imag.*, vol. 10, no. 9, pp. 330–335, Sep. 1991.
[9] A. F. Laine, S. Schuler, J. Fan, and W. Huda, "Mammographic feature enhancement by multiscale feature analisys," *IEEE Trans. Med. Imag.*, vol. 13, no. 4, pp. 725–740, Dec. 1994.
[10] S. Dippel, M. Stahl, R. Wiemker, and T. Blaffer, "Multiscale contrast enhancement for radiographies: Laplacian pyramid versus fast wavelet transform," *IEEE Trans. Med. Imag.*, vol. 21, no. 4, pp. 343–353, Apr. 2002.
[11] J.-L. Starck, F. Murtagh, E. J. Candès, and D. J. Donoho, "Grey and color image contrast enhancement by the curvelet transform," *IEEE Trans. Image Process.*, vol. 12, no. 6, pp. 706–717, Jun. 2003.
[12] S. Mukhopadhyay and B. Chanda, "A multiscale morphological approach to local contrast enhancement," in *Signal Process.*, vol. 80, Apr. 2000, pp. 685–696.
[13] V. Caselles, J.-L. Lisani, J.-M. Morel, and G. Sapiro, "Shape preserving local histogram modification," *IEEE Trans. Image Process.*, vol. 8, pp. 220–230, Feb. 1999.
[14] G. Boccignone, "A multiscale contrast enhancement method," in *Proc. ICIP*, 1997, p. 306.
[15] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," in *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, Jul. 1990, pp. 629–639.
[16] J. A. Stark, "Adaptive image contrast enhancement using generalization of histogram equalization," *IEEE Trans. Image Process.*, vol. 9, no. 5, pp. 889–896, May 2000.
[17] H. Zhu, F. H. Y. Chan, and F. K. Lam, "Image contrast enhancement by constrained local equalization," *Comput. Vis. Image Understand.*, vol. 73, pp. 281–290, Feb. 1999.
[18] H. D.Cheng , M. Xue, and X. J. Shi, "Contrast enhancement based on a novel homogeneity measurement," *Pattern Recognit.*, vol. 36, pp. 2687–2697, Nov. 2003.
[19] N. Petrick, H.-P. Chan, B. Sahiner, and D. Wei, "An adaptive density-weighted contrast enhancement filter for mammographic breast mass detection," *IEEE Trans. Med. Imag.*, vol. 15, no. 2, pp. 59–67, Feb. 1996.
[20] C. M. Bishop, *Neural Networks for Pattern Recognition*. Oxford, U.K.: Clarendon, 1995, pp. 59–73.
[21] G. McLachlan and D. Peel, *Finite Mixture Models*. New York: Wiley, 2000.
[22] A. P. Dhawan, G. Buelloni, and R. Gordon, "Enhancement of mammographic features by optimal adaptive neighborhood image processing," *IEEE Trans. Med. Imag.*, vol. MI-5, pp. 8–15, 1986.
[23] S. Webb, *The Physics of Medical Imaging*. Bristol, U.K.: Adam Hilger, 1988.