

Logarithmic Model-Based Dynamic Range Enhancement of Hip X-Ray Images*

Corneliu Florea, Constantin Vertan, and Laura Florea

Image Processing and Analysis Laboratory, University "Politehnica" of Bucharest,
Romania

Abstract. Digital capture with consumer digital still camera of the radiographic film significantly decreases the dynamic range and, hence, the details visibility. We propose a method that boosts the dynamic range of the processed X-ray image based on the fusion of a set of digital images acquired under different exposure values. The fusion is controlled by a fuzzy-like confidence information and the luminance range is oversampled by using logarithmic image processing operators.

1 Introduction

The X-ray imaging is a widely used technique for medical inspection. Although modern technology provides means and apparatus for digital acquisition, such an option may not be feasible. It is unfortunate, but not always the modern technology has accessible fees. Furthermore, the radiographies acquired with analog means in the past (i.e. film) store valuable information for present medical investigations. Considering the said reasoning, we assumed a low-cost alternative acquisition scheme, which implies photographing the radiographic film with a digital still camera. However, such an approach has a major drawback: the quantity of information available in an radiography is seriously reduced by the low dynamic range of a digital still camera output: the typical radiography produces images that span a dynamic range of some $75dB$, while consumer digital cameras output values in a dynamic range of some $48dB$.

The trivial solution for overcoming the obvious loss of information is to combine frames acquired with different exposures and to posteriorly process the results (involving registration, camera response function (CRF) estimation and frame fusion under various processing models). The resulting quantization oversamples the output space, such that the dynamic range and details visibility are increased. For illustration, we will present examples of high dynamic range images obtained from multiple exposures of a hip prosthesis X-ray. Conclusions and perspectives end the current material.

* This work was supported by the CEEX VIASAN grant 69/2006.

2 Bracketing: Retrieving High Dynamic Range Images from Multiple Exposures

The straightforward solution to the problems generated by the reduced dynamic range of the digital still camera is to combine multiple images of the same scene, taken under various settings (exposure time, aperture). The camera response function (CRF) determines the weights of the mixture parts. This approach is a particular case of super-resolution and is generally known as bracketing. The underlying idea is that each of the images that are to be combined captures with high quality only a certain part of the scene gamut. The bracketing algorithm selects (under the assumption that the multiple images are perfectly aligned), for each pixel of the spatial support scene image, the combination of frames that provides the best value. Thus, an implementation of the dynamic range increase consists of several steps: a first step of image registration (that aligns the multiple images captured from the scene), a step of CRF estimation and the actual image combination (or fusion, or pixel value selection) that computes the enhanced image.

2.1 Image Registration

Image registration means the geometrical alignment of multiple images of a scene based on the matching of the content. Image registration is a widely dealt issue in the field of image processing and several solutions (block matching methods, edges matching methods, object matching methods or global matching methods) are at hand [1].

We used here the robust global matching method of spectrum phase correlation [2], [3]. The underlying idea is based on the translation property of the Fourier transform, F : a translation in the spatial (or time) domain t of a signal x yields a phase shift in the transformed domain.

$$F[x(t + t_0)](\omega) = F[x(t)](\omega) \cdot e^{-j\omega t_0} . \quad (1)$$

Therefore, for a pair of non-aligned images, one will find the corresponding shift as the maximum difference in the phase spectrum of the images. However, the method perform well only if the images exhibit a similar content and if there is no rotational misalignment. The roll component (that produces rotational misalignment) is the least significant motion component for hand-held pictures. If a tripod is considered for capture, imperfections of its mechanical extensions induce only image translations.

2.2 Rough Estimation of Camera Response Function

The CRF (denoted in the current material by g) is the mapping of the device recorded brightness to the scene radiance. The scene radiance is given by the APEX [4] equations as a function of several exposure and device parameters. The

APEX equation that relates the exposure time, the aperture and the incident light is:

$$EV = -\log_2(t) + 2\log_2 N = \frac{S}{K} \int_0^t \phi(t)dt , \quad (2)$$

where EV is the exposure value, the log of t represents the APEX time value (TV), N is the relative diaphragm opening (and log of N represents the APEX aperture value, AV), $\phi(t)$ is the incident light, S is the sensors sensibility (or the amplification for digital cameras) and K is a known constant.

The observation made by Debevec and Malik [5] is of paramount importance for practical bracketing solutions: a set of differently exposed images contains, usually, enough information to recover the CRF using the images themselves.

If the scenario conditions include the same scene, aperture number and amplification as constants, then, by taking into account the right term of equation (2), the measured intensity is linearly dependent of the exposure time. To be more precise, let us assume that images A and B of the same scene were photographed with different exposure times t_A and, respectively, t_B . Given a photo-detector, its charge from the two images must preserve the same ratio as the exposure time. Now, if we come to the reported pixel values u_A and u_B , we get the basic CRF equation:

$$g(u_B) = \frac{t_B}{t_A} g(u_A) . \quad (3)$$

Recovering g from equation (3) is a difficult task ([6]). Certain restrictions have to be imposed on g . The minimum irradiance, 0, will produce no response of the imaging system, hence $g(0) = 0$. The maximum irradiance is an unrecoverable parameter, but the sensors output is limited by a saturation level in the photo-detectors, u_{\max} ; therefore there is an upper-bound : $g(u_{\max}) = D$. The monotonic behavior of g is, also, a typical assumption. Mann and Picard [7] proposed a gamma-like function for g , while Mitsunaga and Nayar [8] used a low degree polynomial regression. Debevec and Malik [5] used a smoothness constraint and recovered the response using a nonparametric model of g , sampled at certain values and represented by a vector.

For our purposes these approaches are too complicated. Further more, it is not feasible to assume that, independently of the frame exposure value, the camera outputs the scene brightness correctly. For over-exposed pictures, it is less likely that pixels having values near the saturation level are accurately recorded. For under-exposed pictures values from the lower part of the range suffer the noise influence and their reported values are corrupted by quantization error. Instead of precise determination of the g function, as in the other mentioned approaches, we will simply compute the confidence that we have in a value recorded at a given exposure bias.

There are different pairs $\{t, N\}$ (exposure time - aperture) that satisfy equation (2). Most of the digital still cameras available on the market are capable of estimating the deviation of the exposure value from the set that balances equation (2). Thus, multiple scenes with the same EV may be obtained; averaging the results will decrease the error of estimation.

Given an exposure value, an image of the usual Macbeth Color Checker chart should exhibit a known set of values. In reality, the camera outputs different brightness intensities. The sum of the squared differences between the output values and the expected values normalized by the expected value is used as an error measure, ε . A low order polynomial regression is implied to extend the domain of the error function from the 24 original values (the number of patches in the chart) to the $[0,255]$ required range. The error function is represented as matrix where the rows are bind to the exposure value parameter, while the columns span the possible gray-levels: $\varepsilon \rightarrow \varepsilon(EV, u)$. The confidence function is computed similarly to a fuzzy negation from the globally normalized error functions:

$$\mu(EV, u) = 1 - \varepsilon(EV, u) , \tag{4}$$

where, again, EV denotes the exposure value and u denotes the gray level.

Examples of non-normalized, interpolated error functions and their corresponding confidence functions computed on images acquired with a SLR-like (Kodak DX6490) digital camera are shown in figure 1.

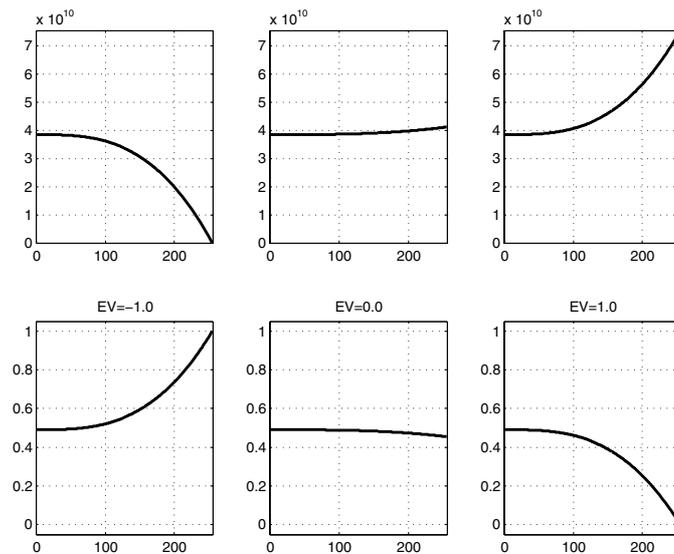


Fig. 1. The top row shows the measured errors with respect to the 0 – 255 gray level range for three exposure values (EV=-1, EV=0, EV=1). The bottom row presents the corresponding confidence functions μ .

2.3 Image Fusion

The image fusion step is the actual dynamic range increasing procedure. A simple approach for fusing a set of N frames taken by a digital camera under several exposures is to discard the pixels with saturated values and to average the remaining values [8]. The frames, denoted by f_1, \dots, f_N , are corrected by the

exposure factor $EV(i)$, such that the pixel located at coordinates (l, m) in the resulting high dynamic range image, f_{HDR} is obtained as:

$$f_{HDR}(l, m) = \frac{1}{N_0} \sum_{i=1}^{N_0} 2^{EV(i)} \cdot f_i(l, m) , \quad (5)$$

where N_0 is the number of frames having non-saturated values at the specified location.

Taking into account the confidence value computed in the previous subsection, a more informative approach is to consider the weighted average (or the convex combination of the pixel values). The weights encode the confidence that a value is outputted correctly. By this approach, the high dynamic range image is computed as:

$$f_{HDR}(l, m) = \frac{\sum_{i=1}^N \mu(EV(i), f_i(l, m)) \cdot 2^{EV(i)} \cdot f_i(l, m)}{\sum_{i=1}^N \mu(EV(i), f_i(l, m))} . \quad (6)$$

3 The Logarithmic Model for Image Fusion

The image values represent, in the case of an X-ray image, the transparency (or the opacity) of the real objects imaged by any given pixel. The underlying physical properties of the imaging system are naturally multiplicative. The key to the logarithmic image processing (LIP) approaches is a homomorphism which transforms the product into a sum (by logarithm), allowing the use of the classical linear filtering in the presence of additive components. Also, it should be clear that the functions used are bounded (taking values in a bounded interval $[0, D)$). During the image processing, the following problem may appear: the mathematical operations on real valued functions use, implicitly, the algebra of the real numbers (i.e. on the whole real axis) and we are faced with results that may fall outside of the interval $[0, D)$ – the physical meaningful values.

3.1 The Classical LIP Model

In the classical LIP model [9], [10], the intensity of an image is completely modelled by its gray tone function v , with $v \in [0, D)$. In this model, the addition of two gray tone functions v_1 and v_2 and multiplication of v by a real number λ are defined in terms of usual \mathbb{R} operations as:

$$v_1 \oplus v_2 = v_1 + v_2 - \frac{v_1 v_2}{D} \quad (7)$$

and respectively:

$$\lambda \otimes v = D - D \left(1 - \frac{v}{D}\right)^\lambda . \quad (8)$$

The use of the operations defined in (7) and (8) leads to an increased visibility of objects in dark areas, as well to the prevention of saturation in high-brightness areas [11].

3.2 The Homomorphic LIP Model

The logarithmic model introduced in [12] works with bounded real sets: the gray-tone values of the involved images, defined in $[0, D)$, is linearly mapped onto the standard set $(-1, 1)$:

$$z = \frac{2}{D} \left(u - \frac{D}{2} \right) \quad (9)$$

where $u \in [0, D)$ and $z \in (-1, 1)$.

The $(-1, 1)$ interval plays the central role in the model: it is endowed with the structure of a linear (moreover: Euclidean) space over the scalar field of real numbers, \mathbb{R} . In this space, the addition between two gray-levels, z_1 and z_2 is defined as:

$$z_1 \oplus z_2 = \frac{z_1 + z_2}{1 + z_1 z_2} \quad (10)$$

while the multiplication of a gray level, z with a real scalar, $\lambda \in \mathbb{R}$ is:

$$\lambda \otimes z = \frac{(1+v)^\lambda - (1-v)^\lambda}{(1+v)^\lambda + (1-v)^\lambda} \cdot \quad (11)$$

3.3 Over-Sampled Fuzed Images

The advantage of the use of LIP models is in the dynamic range reported by the resulting images. If one will examine equation (5) with inputs being all possible combinations of pairs defined between 0 and D , then there will be $2D - 1$ possible resulting levels. If the operation is performed using equation (7), then the number of outputted different levels is in the order of $\frac{D^2}{4}$, while equation (10) leads to an order of $\frac{D^2}{2}$. The logarithmic addition produces an over-sampling of the output values space. The corresponding dynamic range value for $D = 256$ is, roughly:

$$DR = 20 \log \left(\frac{D^2}{2} \right) \approx 90DB \ .$$

Thus, implementing the image fusion in a logarithmic space (or, shortly, by applying log-bracketing) the resulting image will exhibit largely increased number of different brightness levels (which can give the user the possibility of detecting objects in areas displayed uniformly in the original images).

4 Results

The proposed methods were used to enhance hip prostheses X-ray images taken with a consumer digital camera from a original radiographic film placed on a opaque illuminator (negatoscope). For each film a set of images with various exposures (as shown in figure 2) were acquired.

High dynamic range images were produced by the four described approaches: simple averaging (as defined by equation (5)) and CRF weighted averaging (as

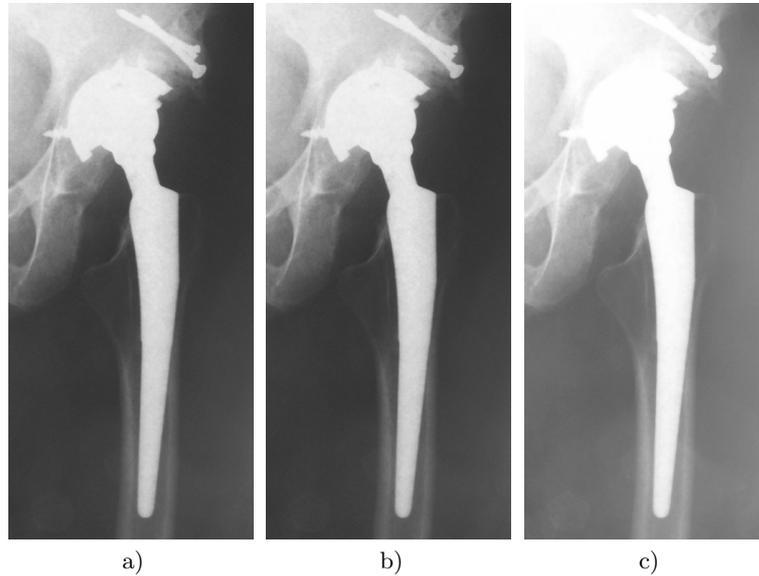


Fig. 2. Originally acquired images: a) under-exposed image ($EV=-1$); b) correctly exposed image ($EV=0$); c) over-exposed image ($EV=1$)

defined by equation (6)) implemented with classical \mathbb{R} addition/ multiplication and with LIP model (both classical and homomorphic) addition/ multiplication. The intensity values were quantized with 12 bits per pixel (bpp) precision. Figure 3 presents an example of such high dynamic range X-ray images. The 12 bpp gray level images were displayed on usual RGB color displays using an extension of the classical gray level map via highly unsaturated colors that match the needed luminance levels that uniformly sample the luminance range. Indeed, the human visual system is unable to distinguish colors for which the difference within the maximal and minimal RGB components is small (less than 5 units on the 256 units scale). As such, the 4096 gray levels needed for the 12 bpp representation are obtained from the 256 classical (and exact) gray levels and 3840 highly unsaturated colors.

The criteria used for choosing the best picture is the number of visible details of the prosthesis and the distinction between its parts, the visibility of the bone channel surrounding the prosthesis tail and the visibility of the bone fibres structure. Under such criteria, the high dynamic range images computed using the convex combination are the best. The direct implementation, in this case, leads to several outcomes, like the smearing effect on the background (which is expected to be completely dark) or less contrast in the prosthesis tail area. The images computed using the convex combination implemented according to the LIP model are the best. Figure 4 shows some of the relevant prosthesis details.

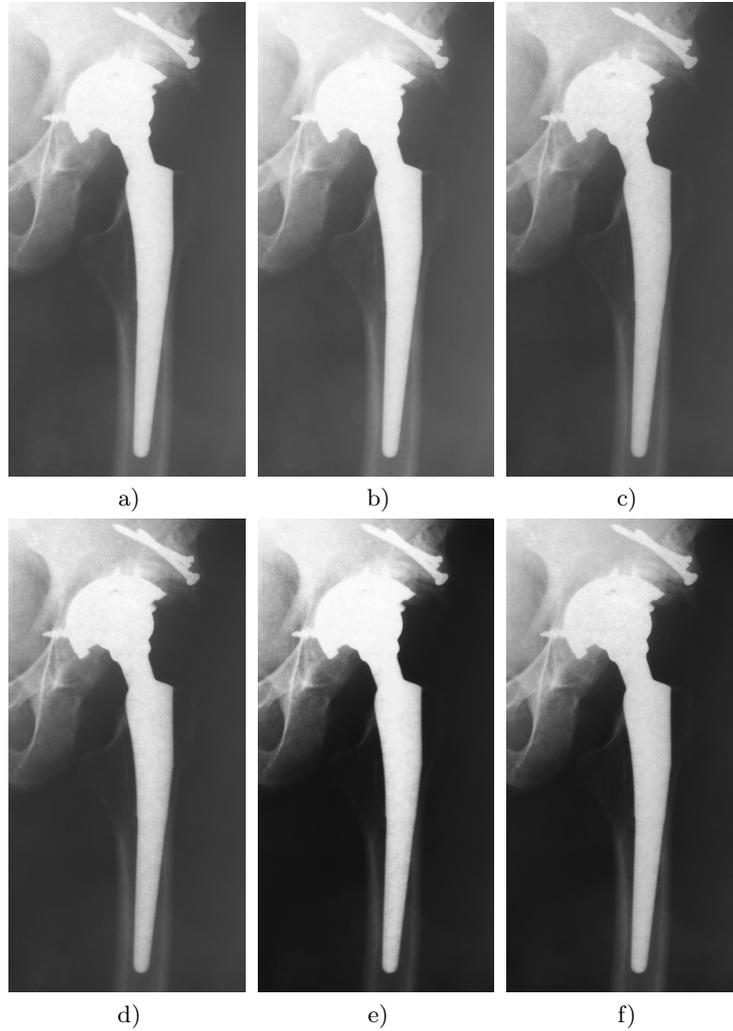


Fig. 3. High dynamic range images obtained from the set presented in figure 2 by averaging (as defined by equation (5)) using a) IR addition and multiplication, b) classical LIP addition and multiplication c) homomorphic LIP addition and multiplication and by CRF weighted averaging (as defined by equation (6)) using d) IR addition and multiplication, e) classical LIP addition and multiplication f) homomorphic LIP addition and multiplication

5 Conclusions

We presented a new method that takes as input a set of X-ray frame-images with the same subject, but different exposure values and combines them into a high-dynamic range image. The proposed fusion scheme requires confidence

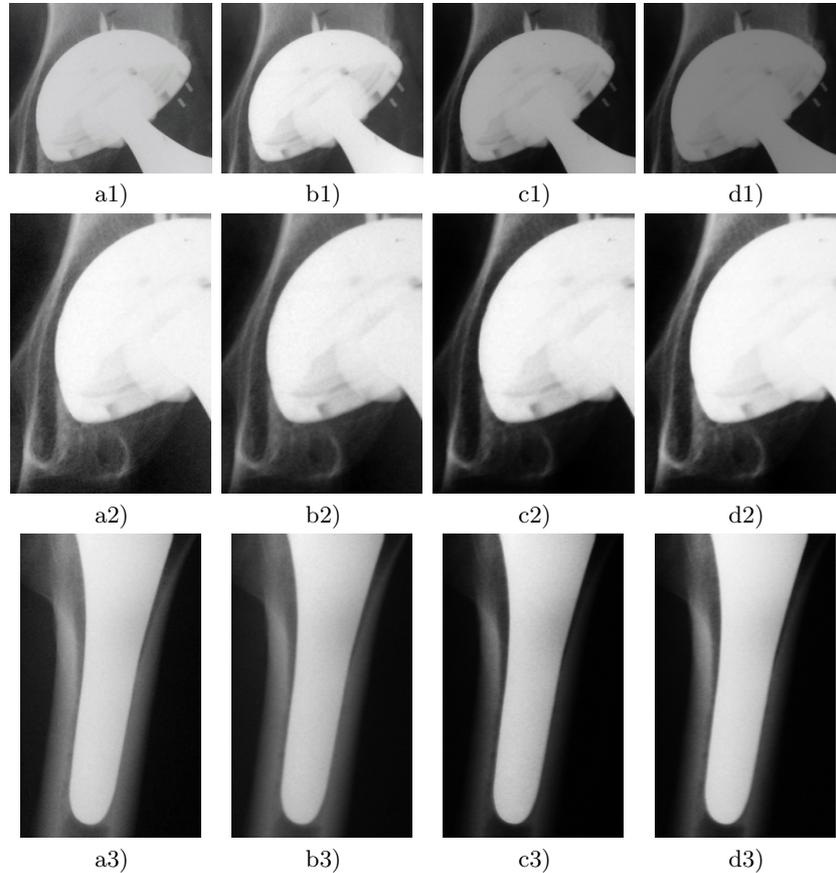


Fig. 4. Details from X-ray prosthesis images: top two rows – prosthesis head and cup, bottom row – prosthesis tail. The images are: a) well exposed original images ($EV=0$) and high dynamic range images obtained by CRF weighted averaging (as defined by equation (6)) using b) \mathbb{R} addition and multiplication, b) classical LIP addition and multiplication d) homomorphic LIP addition and multiplication. The classical LIP model seems to yield the greatest detail visibility.

information derived from the non-linearity of the camera response function. Performing the operation required by the fusion scheme according to a logarithmic image processing method highly increases the number of resulting gray levels. Therefore objects placed in uniform areas become easier to examine. The proposed method was successfully applied to enhance the dynamic range of hip prosthesis X-ray film images acquired by a consumer digital camera. Even that the classical LIP model was designed to be used for special categories of images, there are proves that the homeomorphic LIP model is suitable for most of the images. By these means we intend to test the currently described method on natural images.

References

1. Schechner, Y.Y., Nayar, S.K.: Generalized mosaicing: High dynamic range in a wide field of view. *International Journal on Computer Vision* 53, 245–267 (2003)
2. Kuglin, C.D., Hines, D.C.: The phase correlation image alignment method. In: *Proc. of IEEE Conference on Cybernetics and Society*, Bucharest, Romania, pp. 163–165. IEEE Computer Society Press, Los Alamitos (1975)
3. Averbuch, A., Keller, Y.: Fft based image registration. In: *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing ICASSP '02*, Orlando FL, USA, vol. 4, pp. 3608–3611. IEEE, Los Alamitos (2002)
4. PH2.5-1960, A.: American standard method for determining speed of photographic negative materials (monochrome, continuous tone) United States of America Standards Institute (1960)
5. Debevec, P., Malik, J.: Recovering high dynamic range radiance maps from photographs. In: *Proc. of ACM SIGGRAPH 24th Annual Conference on Computer Graphics and Interactive Techniques*, Los Angeles CA, USA, vol. 1, pp. 369–378. ACM, New York (1997)
6. Grossberg, M.D., Nayar, S.K.: High dynamic range from multiple images: Which exposures to combine? In: *Proc. of IEEE Workshop on Color and Photometric Methods in Computer Vision at ICCV 2003*, Nice, France, IEEE, Los Alamitos (2003)
7. Mann, S., Picard, R.: Being 'undigital' with digital cameras: Extending dynamic range by combining differently exposed pictures. In: *Proc. of ST's 48th Annual Conference*, Washington D.C. USA, vol. 1, pp. 422–428 (1995)
8. Mitsunaga, T., Nayar, S.K.: High dynamic range imaging: Spatially varying pixel exposures. In: *Proc. of IEEE Conference on Computer Vision and Pattern Recognition CVPR*, Hilton Head SC, USA, vol. 1, pp. 472–479. IEEE, Los Alamitos (2000)
9. Jourlin, M., Pinoli, J.C.: A model for logarithmic image processing. *Journal of Microscopy* 149, 21–35 (1998)
10. Jourlin, M., Pinoli, J.C.: Logarithmic image processing. *Advances in Imaging and Electron Physics* 115, 129–196 (2001)
11. Deng, G., Cahill, L.W., Tobin, G.R.: The study of logarithmic image processing model and its application to image enhancement. *IEEE Trans. on Image Processing* 4, 506–512 (1995)
12. Patrascu, V., Buzuloiu, V., Vertan, C.: Fuzzy image enhancement in the framework of logarithmic model. In: *Nachtegael, M., Kerre, E. (eds.) Algorithms in Modern Mathematics and Computer Science. Studies in Fuzziness and Soft Computing*, vol. 122, pp. 219–237. Springer Verlag, Heidelberg (2003)