DYNAMIC RANGE ENHANCEMENT OF CONSUMER DIGITAL CAMERA ACQUIRED HIP PROSTHESIS X-RAY IMAGES

Laura Florea, Constantin Vertan, Corneliu Florea, Alina Oprea

Image Processing and Analysis Laboratory, Politehnica University of Bucharest
{lsirbu,cvertan,cflorea,aoprea}@alpha.imag.pub.ro

ABSTRACT

The X-ray film image is the main medical diagnosis tool in the evaluation of the fit of the hip prostheses inserted in total hip arthroplasty (THA) procedures. In the design of a computer-aided diagnosis tool, one of the most important operations is the digital capture of the film images. This contribution investigates the use of a consumer-grade digital still camera as a digital acquisition tool for hip prosthesis X-ray images. We propose and discuss two methods for the increase of the dynamic range and contrast of the digital image, using single or multiple pictures of the radiological film. The prosthesis stem is segmented from the resulting digital images enabling the measurement of basic clinical parameters (stem fit within the medular channel, Gruen areas identification).

1. INTRODUCTION

The idea of total hip prosthesis was born before the 1950s and evolved towards the nowadays total hip prosthesis with a stem and an acetabular component. Total hip arthroplasty (THA) is, at present, a well-known and a highly developed technique to reduce pain in arthrotic and arthritic hips; in the US alone more than 150000 THAs are performed every year [1]. One of the latest developments in total hip arthroplasty is the usage of uncemented prosthesis with hydroxylapatite coating. The hip protheses require a continuous, regular and careful follow-up and monitoring, in order to detect failures. Wear is an important factor in failure of prostheses, as is aseptic loosening, which is an indication for revision surgery in up to 20 % of the primary total hip arthroplasties. Aseptic loosening of a total hip prosthesis is the result of a combination of bio-mechanical and chemical factors: the weakening of the bone resulting from bone resorption because of particle disease; the material strain in the interface of prosthesis and bone; the failure of the ingrowth in improperly fitted uncemented prostheses; the inadequate stress transfer of prosthesis to bone [2].

A common characteristic of the hip prosthesis scoring (or fit evaluation) systems is the combination of clinical, patient-subjective scores (as the Harris hip score or Oxford score) with radiological gradings (obtained by subjective orthopedic evaluation) and bone measurements (such as the bone mass density). The minimal yearly investigation requires the use of two X-ray images (frontal and lateral) of the hip area.

A key issue in the prosthesis check-up is the need of consulting very specialized medical personnel, which is concentrated in a few centers, unlike the geographical spread of X-ray machines and general practitioners. Under these circumstances, the implementation of a computer-assisted diagnosis system is of great interest, since it may help distribute some of the regular, normally evolving cases away from the THA-performing medical centers and orthopedic surgeons.

The very first step in the implementation of such a computer-assisted diagnosis system is the digital acquisition of common, film X-ray images with generally available, unexpensive digitization equipment. The recent advances in consumer electronics made the digital still camera the first choice for such a task. Digitization of an X-ray film will simply mean to take photos of the negatoscope-presented film. Yet, the matter is not simple, since there is an obvious loss of image quality in the digital acquisition of an X-ray film by digital photography. The spatial resolution of an X-ray is given by the grain density of the film (typically at some $10^8 grains/cm^2$); the dynamic range of the radiologic film is approximately 75dB [3]. The images acquired by a digital still camera have a spatial resolution given by the camera sensor size (upper limited at some 10^7 pixels) and the size of the imaged area. Such images exhibit a typical dynamic range of some 48dB to 72dB (depending on the image mode - color or RAW).

This paper proposes two approaches to the dynamic range enhancement for digitized images of the hip prosthesis X-ray and their joint use in an automatical feature-space segmentation of the image, that can reliably detect the prosthesis and the femoral bone. Several prosthesis to bone distance measurements are performed and an overall prosthesis fixation index is computed. The remainder of the paper is organized as follows: the next section presents the two approaches to the dynamic range increase: classical bracketing and pixel ordering. Section 3 describes the prosthesis and bone segmentation procedure, respectively the clinical relevant measurements and section 4 presents some conclusions and directions of further investigation.

2. X-RAY IMAGE DYNAMIC RANGE INCREASE

Studying the behavior of the femoral components of the hip prosthesis can be limited to the imaging of the immediate area surrounding the prosthesis stem, that is an area of some 300 cm^2 of radiographic film. From the medical practice it comes that we should reliably measure and represent details as small as $0.1mm$ within the area of interest. That leads to the constraint of using a minimal 10 pixels/mm (250ppi) resolution, implying the use of a digital still camera exhibiting a sensor of at least 4 Mpixels. This is easily achievable by a reasonable consumer digital still camera; yet the camera cannot directly provide the equivalent dynamic range of the X-ray film.

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

2.1 X-ray image dynamic range increase by multiple scene instances

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 (shutter speed, f-number). This approach is part of the super-resolution algorithms 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 color (grayscale) gamut (as imposed by the settings of the camera). The bracketing algorithm selects, for each pixel of the spatial support of the scene the acquired image that provides for it the best value. That selection is based on the assumption that the multiple images are perfectly aligned.

Thus, our implementation of dynamic range increase based on multiple scene instances works in two steps: a first step of image registration (that aligns the multiple images captured from the scene) and the actual image combination (or fusion, or pixel value selection), that computes the enhanced image.

2.1.1 Image registration

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

We used the robust global matching method of spectrum phase correlation [4], [5]. The underlying idea is based on the translation property of the Fourier transform F: a translation in the spatial (or time) domain x of the signal s leads to phase shift in the transformed domain.

$$F\left[s(x+x_0)\right] = F\left[s(x)\right] e^{-j\omega x_0}. \quad (1)$$

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 is designed to work if the images are similar by content and if there is no rotational misalignment, which is the case in our acquisition setup.

2.1.2 Image fusion

This second step is the actual dynamic range increasing. The approach is widely described in the literature [6], [7], [8], [9]. In general, the proposed method is, in some extent, similar to the one proposed in [9]. However, we exploit the particular characteristics of the images generated from hip prosthesis X-rays.

The dynamic range of an image is given by:

$$DR = 20\log\left(\frac{I_{\max}}{I_{\min}} \cdot \frac{e_{\max}}{e_{\min}}\right)[dB], \quad (2)$$

where $I_{\max}$ and $I_{\min}$ are the maximal and minimal representable intensities and $e_{\max}$ and $e_{\min}$ are the maximal and minimal exposure values.

The exposure of a picture, using the standard APEX system, is given by:

$$EV = TV + AV \quad (3)$$

where, $TV = -\log_2 t$ and respectively $AV = 2\log_2 N$. Here N is the relative aperture (f-number) and t is the exposure time (shutter speed). The exposure value is related to the scene illumination by:

$$EV = \frac{L \cdot S}{K}, \quad (4)$$

where L is the average scene illumination, S is the sensor speed (usually called ISO parameter) and K is a constant.

Therefore, changes of the exposure value will modify the captured incident illumination. A highly exposed image will be saturated in the bright scene areas, but will capture dark regions well. In contrast, an under exposed image will have less saturation in bright regions, but will end up being too dark and noisy in the dark areas. The complementary nature of these images allows one to combine them into a single high dynamic range image. Variation of the exposure value may be obtained by changing the shutter speed or by increasing and decreasing the aperture.

In our setup we modified the exposure time. A set of such pictures may be seen in figure 1. If, for instance we will choose, in equation 2, $e_{max} = 32e_{min}$, the dynamic range of the resulting image will be $78.2dB$. Such a value is obtained after combining 6 pictures with e_{max}=e_1=$2e_2$=$4e_3$=$8e_4$=$16e_5$=$32e_6$=$32e_{min}$.

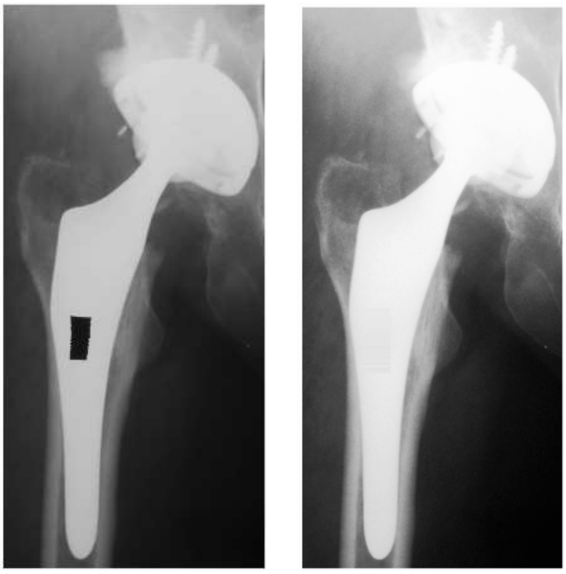
a) b) c)

Figure 1: Set of photographic images of a hip prosthesis X-ray taken under various exposures; from left to right exposures e_6, e_4 and e_2.

The fusion method is direct. The saturated pixels and the noise corrupted pixels are discarded, and we simply average the values of the remaining pixels across the set of acquired images. The pixels arriving from less exposed images add details. The picture with large exposure provides the base of the image content. The operation is performed independently for each pixel in the resulting image. The final dynamic range enhanced image is shown in figure 2 a).

2.2 X-ray image dynamic range increase by total pixel ordering

The bracketing method described in the subsection above has the disadvantage of requiring several frames which must be properly registered. It is not a rare case that image registration fails. The natural way to avoid such a problem is the use of a single image, reasonably well exposed, and to review the dynamic range increase problem within an image enhancement framework.



a) b)

Figure 2: Dynamic range enhanced images by the presented approaches: a) Bracketing; b) Pixel ordering based on the proposed features (7).

Basically, we can view the dynamic range increase as the creation of significantly more different pixel values that the number of different pixel values available in the initial image. This cannot be performed by simple, point image enhancement operations, such as the classical histogram equalization or histogram stretching [10], but requires the segregation of same-valued pixels according to their local statistics. In some way, we replace the multiple luminance values available at each pixel location from multiples acquisitions of the scene (as used in the bracketing approach) by data (descriptors) computed in some spatial neighborhoods, yielding to the association of a feature vector to each image pixel. The pixels are then ordered according to their feature vectors and from the resulting string pixels are assigned new, increasing gray level values. This is the approach introduced in [11] for histogram specification.

Usually, the specified histogram is either flat (yielding to a perfect histogram equalization) or a stretched variant of the original histogram of the image, but with the same number of bins, corresponding to the number of gray quantization levels. In the following, we will use a different approach, by specifying a histogram with more bins (and, thus, more different gray levels in the resulting image) than the number of bins in the original image.

Let I be the initial image. Each pixel (i, j) from the image I is mapped to a feature vector $\mathbf{g}(i, j)$, composed by:

$$\mathbf{g}(i,j) = [g_1(i,j), g_2(i,j), g_3(i,j), \ldots] \quad (5)$$

where $g_1(i, j)$ is the initial pixel value (in order to maintain the relative intensity ranking of pixels) and $g_k(i, j)$ with $k > 1$ are features computed in some neighborhoods of the current pixel. Let $\prec$ be the lexicographic ordering relation applied to the feature vectors $\mathbf{g}(i, j)$; two feature vectors verify the ordering relation $\mathbf{g}(i_1, j_1) \prec \mathbf{g}(i_2, j_2)$ if and only if:

$$\exists k, \begin{cases} g_l(i_1,j_1) = g_l(i_2,j_2), & 1 \leq l < k \\ g_k(i_1,j_1) < g_k(i_2,j_2). \end{cases} \quad (6)$$

The choice of specific features $g_k(i, j)$ is imposed by the desired appearance of the final image. The use of local averages, as suggested in [11], as a reference choice, will favor the pixels located in uniform areas and will slightly blur the edges. In order to enhance the overall edge visibility, we shall use some different form for the $g_k(i, j)$ features, as suggested by the following observations.

Considering the case of two pixels that have the same value, their difference is imposed by the relative order of their features. If the pixels are located in an uniform region, then their supplementary features should be zero. If the pixels are located within a non-uniform area, then the contrast of that area should be increased. It is well known that the human eye is sensitive to luminance contrast and two regions of the same shape and gray level will be perceived differently based on the gray level of the background. If one of the shapes is placed on a lighter background (higher gray level), it will appear darker than the other one. Thus, in the non-uniformity areas, the pixel that has a lower gray level than the local mean will be made even darker; the pixel that has a greater gray level compared to the local mean will be made lighter.

We define the features $g_k(i, j)$ in $(2k+1) \times (2k+1)$ neighborhood as follows:

$$g_k(i,j) = \alpha(\sigma) sign(I(i,j) - \overline{I_k(i,j)}), \quad (7)$$

where $\overline{I_k(i,j)}$ is the local mean (within the $(2k+1) \times (2k+1)$ neighborhood) and α is a function of the local standard deviation σ. An intuitive choice for α may be:

$$\alpha(\sigma_k) = \frac{\sigma_k}{\sigma_{\max}}, \quad (8)$$

where $\sigma_{\max}$ is the maximal standard deviation within the image (computed at each pixel location within the $(2k+1) \times (2k+1)$ neighborhood).

The described method was applied using a three-dimensional feature vector on the X-ray image captured under an automatic exposure (as the one presented in figure 1 b)). The result of the method is shown in figure 2 b).

3. X-RAY IMAGE SEGMENTATION AND PROSTHESIS FIT ANALYSIS

As previously explained, the final goal of the implemented system is the automated analysis of the elements of interest from the hip prosthesis X-ray: the prosthesis and the femoral bone. The femoral part of the uncemented hip prosthesis is a metallic component, that spans the most bright gray levels in the digitally acquired X-ray image and can be identified as the rightmost mode of the image histogram. The neighboring femoral bone and the soft tissues and air contribute with two partially overlapped modes, as shown by the histogram in figure 3.

Since the modes that represent the components of interest are overlapped, simple histogram thresholding [10] is not an effective solution. We used a classical fuzzy C-means (FCM) clustering algorithm [12] for the identification of the three classes within the image. Figure 4 shows the image of the membership degree of the bone class and the segmentation that results from a simple, fixed threshold defuzzification.

Once the prosthesis is segmented, the limits of the Gruen zones [1], [2], [13] are obtained from simple geometrical

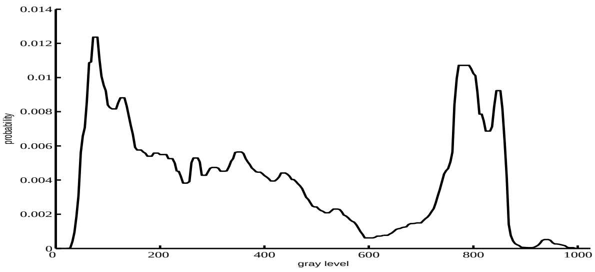
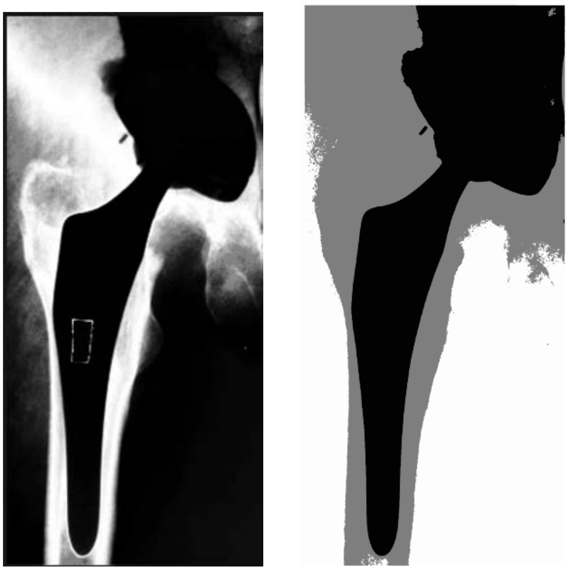


Figure 3: Typical histogram of a digital X-ray image of an uncemented hip prosthesis (we shall note that the gray levels are coded on 10 bits, as a result of the dynamic range enhancement).



a) b)

Figure 4: a) Fuzzy membership degree within the bone class; b) Crude 3-class segmentation of the prosthesis image using the increased dynamic range images from figure 2 a).

considerations and for each zone the distance from the prosthesis to the cortical wall is computed. The typical profile of the distance variation around the prosthesis is shown in figure 5, showing the fixation points of the prosthesis stem (the near-zero minima in the distance plot).

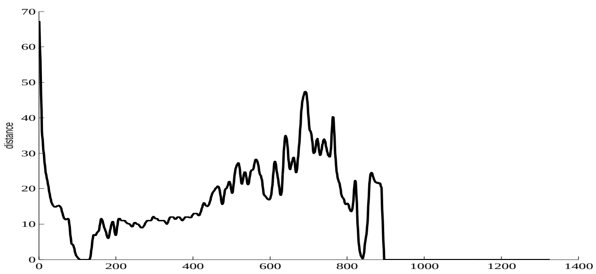


Figure 5: Typical distance variation around the prosthesis for Gruen areas 5, 6 and 7 (right-hand contour of the prosthesis shown in figure 4).

4. CONCLUSIONS

This paper presented a new dynamic range increase method based on pixel ordering that uses a single input image (opposed to the multiple images used by the bracketing algorithm). The method provides good results and is used in the framework of cheap digitization of X-ray films by the use of consumer digital still cameras. Simple segmentation tools can be effectively used on the produced digital X-ray image in order to identify and characterize elements of interest in the particular area of orthopedics - hip prosthesis monitoring.

REFERENCES

[1] A. G. Della Valle and W. G. Paprosky, ”Revision total hip arthroplasty”, Orthopaedic Knowledge Online, American Academy of Orthopaedic Surgeons, 2003.

[2] J. D’Antonio, J. C. McCarthy, W. L. Barger et al, ”Classification of femoral abnormalities in total hip arthroplasty”, Clinical Orthopedy, vol. 296, pp. 133-139, 1993.

[3] S. Webb: The physics of medical imaging. Institute of Physics Publishing, 1998.

[4] C. D. Kuglin and D. C. Hines, ”The phase correlation image alignment method”, in Proc. of IEEE Conference on Cybernetics and Society, Sept. 1975, pp. 163165.

[5] A. Averbuch and Y. Keller, ”FFT based image registration”, in Proc. of IEEE ICASSP ”02, 2002, vol. 4, pp. 3608-3611.

[6] P.J. Burt and, R.J. Kolczynski, ”Enhanced image capture through fusion”, in Proc. of ICCV, 11-14 May 1993, pp 173-182.

[7] M. D. Grossberg and S. K. Nayar, ”High Dynamic Range from Multiple Images: Which Exposures to Combine?”, in Proc. of ICCV 2003 Workshop on Color and Photometric Methods in Computer Vision.

[8] S. K. Nayar and T. Mitsunaga, ”High Dynamic Range Imaging: Spatially Varying Pixel Exposures”, in Proc. of IEEE CVPR, Hilton Head Island, South Carolina, June, 2000.

[9] J. Yang and R. S. Blum, ”A Region-Based Image Fusion Method Using the Expectation-Maximization Algorithm”, in Proc. of Conference on Information Science and Systems, Princeton, NJ, March 2006.

[10] R. C. Gonzales and R. E. Woods, *Digital Image Processing*, Addison-Wesley, 1992.

[11] D. Colţuc and P. Bolon, Chossery J.M. ”Exact Histogram Specification”, in IEEE Trans. on Image Processing, vol. 15, no. 2, pp 1143-1152, 2006.

[12] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Functions*, Plenum Press, New York, 1981.

[13] M. R. Downing, P. E. Undrill, P. Ashcroft, D. W. L. Hukins, and J. D. Hutchison, ”Automated femoral measurement in total hip replacement radiographs”, in Proc. IEE Image Processing and its Applications, 14-17 July 1997, vol. 2, pp. 843 - 847, 1997.