

# STATISTICAL ANALYSIS OF ULTRASOUND ECHO FOR SKIN LESIONS CLASSIFICATION

HANNA PIOTRZKOWSKA, JERZY LITNIEWSKI, ELŻBIETA SZYMAŃSKA\*, ANDRZEJ NOWICKI

Institute of Fundamental Technological Research, Department of Ultrasound  
ul. Pawińskiego 5b, 02-106 Warsaw, Poland  
e-mail: hpiotrzk@ippt.gov.pl

\*CSK MSWiA Hospital, Dermatology Clinic  
ul. Wołoska 137, 02-570 Warsaw, Poland

*Propagation of ultrasonic waves in the tissue is sensitive to the alternation of tissue composition and structure.. This paper presents the classification of healthy skin and skin lesions (basal cell carcinoma (BCC)) based on statistic parameters of the envelope of echosignal. The statistics of envelope of the ultrasonic signal was modeled using Rayleigh and non-Rayleigh (the K-distribution) statistics. Furthermore the characteristic parameter of K-distribution, the effective number of scatterers ( $M$ ) was investigated.*

*Comparison of the results obtained for region of the skin where the BCC was diagnosed and the regions of healthy skin has shown differences in the values of  $M$  parameter. These results indicate that this parameter has the potential for extracting information useful for characterizing skin lesions.*

## INTRODUCTION

The ultrasonic echo-signal received by the transducer is formed from the summation of partial echoes, , from the scatterers located in the tissue. The interference of waves, reflected from the randomly located scatterers, results in speckles formation, which in ultrasound images, can be seen as random process whose statistical features depend on the tissue class. The character of the echo-pulse depends on the number of scatterers in the resolution cell and their size comparing to wavelength of ultrasound signal. When the number of partial echoes, and scatterers contained in resolution cell and involved in signal formation, is large, speckles intensity (signal envelope) is Rayleigh distributed. In a case when scatterers number is low the Rayleigh distribution is not applicable. In this case envelope from tissues may be modeled in terms of non-Rayleigh statistics.

In cases when a disease changes the properties of tissue scattering, the ultrasonic echo provides potentially useful information for detection and diagnosis. In this work the expected

differences in scatterer properties between the healthy skin and skin lesions were the motivation to examine the correlation of texture and signal parameters for healthy skin and the skin with basal cell carcinoma. Our research was concentrated on checking possibilities of K distribution as a model suitable for modeling statistics envelope of signal obtained from human dermis in vivo. This paper presents the classification of healthy skin and skin with diagnosed basal cell carcinoma based on characteristic parameter of K distribution the effective number of scatterers.

The basal cell carcinoma (BCC) is one of the most common skin cancer. Tumor is composed of uniform cells that appear like basal cells with dark nuclei. The cells lie in groups or nets and are surrounded by dermal stroma.

Figure 1 presents histological picture of healthy skin (a) and histological picture of basal cell carcinoma (b) [9].

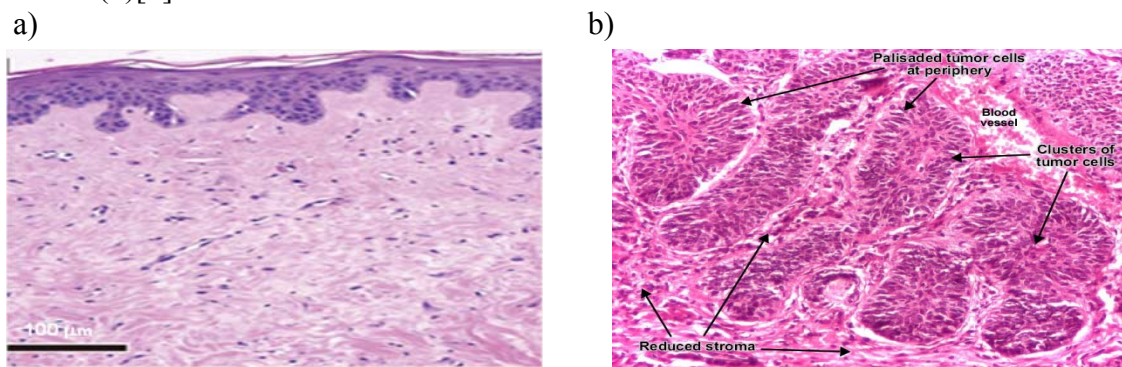


Fig.1. The histological picture of healthy skin (a) and skin with diagnosed basal cell carcinoma (b).

## 1. MATERIALS AND METHODS

### 1.1 MODEL FOR SCATTERING

When the tissue is interrogated with ultrasounds the signal, representing the acoustic pressure received at the transducer, is a sum of acoustic pressures, reflected by large number of scattering points, located within a resolution cell. The resolution cell is defined by the pulse length and the area of the beam cross-section.

The ultrasonic scattering from tissues has generally been modeled using Gaussian model, which assumes that there is a large number of scatterers having identical scattering cross-sections, randomly located in resolution cell. Under these conditions, the probability density function of envelope of the echo will be Rayleigh distributed.

In some cases echoes are not well represented by the Gaussian speckle model. The non-Gaussian statistics for echo data can result at two situations: too few scatterers per resolution cell volume to provide the uniform-phase condition required for fully developed speckle and variation in scattering intensity from each scatterer. In this situation the probability density function of envelope will not remain Rayleigh distributed.

### 1.2 STATISTICS OF THE ECHO-ENVELOPE

The echo signal from scattering medium, such as biological tissue, can be modeled as the sum of individual backscattered signals from a number of scattering points within the medium. Backscattered signal of each individual scatterer can be represented as a phasor. When amplitude and phase are the random variables due to properties and scatterers position, the sum is also random. In this case, tip of phasor representing echo signal, is found at the

“random walk”, corresponding to the end-to-end sum of  $N$  random phasors  $\mathbf{s}$ . Using the phasor notation [1],[2] echo signal can be described as resultant random phasor sum:

$$\mathbf{s} = S e^{j\theta} = \sum_{i=1}^N \mathbf{s}_i = \sum_{i=1}^N s_i e^{j\theta_i} \quad (1)$$

where  $N$  means number of phasor components,  $S$  (a complex number) represents the resultant phasor,  $S$  describes amplitude and  $\theta$  describes phase of the resultant. The  $\mathbf{s}_i$  represents  $i$ th component of phasor in the sum ( $s_i$  and  $\theta_i$  are the amplitude and phase  $s_i$  respectively).

If amplitudes  $s_i$  and phases  $\theta_i$  are statistically independent and uniformly distributed and the  $N \rightarrow \infty$ , it can be show that probability density function of the envelope is given by:

$$p(S) = \frac{S}{\sigma^2} e^{-\left(\frac{S^2}{2\sigma^2}\right)} \quad \sigma > 0 \quad (2)$$

The result is known as Rayleigh density function, where  $\sigma$  means scaling parameter .

One interesting property of this distribution is the value signal-to-noise ratio, which is defined to be ratio of mean to standard deviation (MSD):

$$\text{MSD} = \frac{E[S]}{[E[S^2]-E[S]^2]^{1/2}} = 1.913 \quad (3)$$

and is constant and independent of the backscattered energy. This MSD level can be used to detect to “fully formed” speckle, any departure of MSD from the value of 1.913 can be regarded as the departure from the Rayleigh distribution.

The Rayleigh distribution model suits ideal ultrasound backscatter conditions. Unfortunately in typical applications of medical ultrasound, the backscatter signal cannot be assumed to the Rayleigh distributed.

In many cases, following conditions leads to necessity applying another kind of distribution: the number of scatterers per resolution cell might not be large enough or the scatterers might not be located randomly (due to periodicity, clusterings).

It is worth noting that there is no one to one relationship between the number of scatterers and the number of scattering sites. The scattering sites are clustered at certain areas, which depends on the scatterer size and arrangement. Degree of the clustering, not the number of scattering sites has a close relationship with the scatterers concentration.

One of the non-Rayleigh distributions which can be use to describe statistics of envelope of signal is so-called K distribution [4].

The K distribution can represent the statistics of the echosignal from tissues under varying conditions. The K distribution is appropriate for a region where there are a large number of randomly distributed scatterers having fairly uniform scattering cross-section (Rayleigh), scatterers having widely varying scattering cross-sections or small number of scatterers and even scatterers having some sort of periodic spacings [3].

The K distribution is given by probably destiny function:

$$p(A) = 2 \left(\frac{A}{2}\right)^M \frac{b^{M+1}}{\Gamma(M)} K_{M-1}(bA) \quad A \geq 0; \quad M, b > 0 \quad (4)$$

where  $b = \sqrt{\frac{4M}{E[S^2]}}$  and  $K_\beta(\cdot)$  is modified Bessel function of the second kind of order  $\beta$ ,  $\Gamma(\cdot)$  is the standard Gamma function.

Jakeman and Pusey [3] have shown that the parameter  $M$ , which can be treated as the “effective” number of scatterers per resolution cell, is given as:

$$M = N_s(1 + v) \quad v > -1 \quad (5)$$

where  $N_s$  is the number of scatterers in resolution cell of echo imaging system, the parameter  $\nu$  modulates the number of scatterers to give the „effective” numbers of scatterers affecting at the echo envelope statistics.

One of the various M parameter estimation method, is the “method of moments” proposed by Jakeman and Pusey [3] and Weng and *al.*[9]. In this work second and fourth order moment method was used. The fourth normalized central moment is given by:

$$r_4 = \frac{E[S^4]}{[E[S^2]]^2} = 2 \left( 1 + \frac{1}{M} \right) \quad (6)$$

so

$$M = \frac{2}{r_4 - 2} = \frac{2}{\frac{E[S^4]}{[E[S^2]]^2} - 2} \quad (7)$$

In equation (5) it is easy to see that “effective” number of scatterers estimated by the K distribution model depends on the actual number of scattering sites per resolution cell as well as the statistics of backscatter coefficient. The type and distribution of scatterers are intrinsically related to the type of tissues that the ultrasound beam is passing through. Therefore the M parameter can be used to distinguish between regions differing in special density of scatterer or between regions of varying scatterer’s cross-section and in consequence can be used as a parameter for tissue characterization.

By comparing the statistics of the *in vivo* data to the K distribution, as well as Rayleigh distribution, the validity of both distributions as an accurate descriptors of the envelope of echo was established in this work.

The goodness of the fit of K-distribution and Rayleigh distribution to the empirical histograms was evaluated using the mean square error (MSE). The small value of MSE indicates the good fit of the distribution to the empirical data.

## 2. MEASUREMENT PROCEDURE

Skin tissues were examined *in vivo*. The measurements were performed in the dermis at the various parts of the body. Skin tissues were examined *in vivo*. Two kinds of data was obtained for patients of Dermatology Clinic. The first measurements were done in region of skin where the basal cell carcinoma was diagnosed and the second in the fragments of skin with no pathological changes.

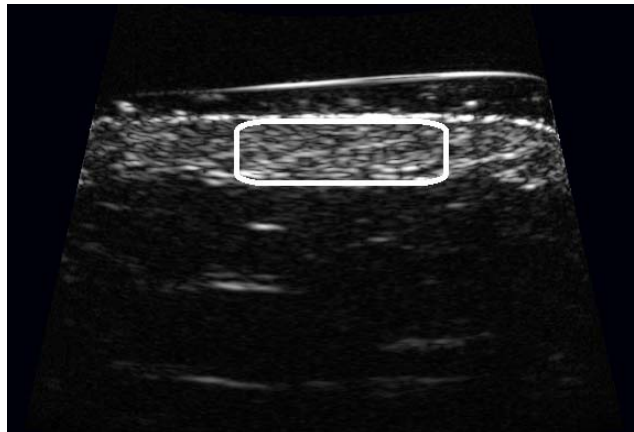


Fig.2. The B-mode image of healthy human skin with marked ROI.

For computing the above echo statistics parameters, the region of interest (ROI) corresponding the BSS was chosen for pathological skin and for the healthy skin the ROI was placed in the region of dermis, where on the B-mode image, the tissue seems to be the most homogeneous.

The skin scanner, which was used for the data from the skin acquisition, worked at frequency 25 MHz. This scanner was developed in our laboratory [5]. It performed a sector scan with the image frame rate up to 10 Hz. The transmitted signal and scattered echoes were sampled at 200 MHz frequency with 12 bits resolution. In this study we have used a 20  $\mu\text{m}$  thick spherical transducer (with 3 mm diameter, 8.6 mm focal length) made of the modified PZT 37 deposited on the PZT substratum using the thick-film technology (Ferroperm, Denmark). The received sequences were envelope detected and displayed. Simultaneously, the RF data were stored separately.

### 3. SIGNAL PROCESSING

Prior to the statistical evaluation of the received RF signals it was necessary to compensate the signal for the formerly induced Time Gain Control (TGC) and in the next step the preprocessing procedures echo data was compensated for the attenuation in the tissue.

The attenuation coefficient  $\alpha(f)$  of the skin was determined for every case separately, using the spectral difference technique based on the comparison of the power spectra of the signals backscattered at different depth in the tissue [7].

For the compensation of attenuation the following algorithm was used [6]. First, the spectrum of attenuated signal (FA) was calculated. Next, the synthesis of a new signal ( $F(\cdot)$ ) on the basis of spectral components of the backscattered signal was performed. During the synthesis, the amplitudes of spectral components were increasing with the increasing value of the depth co-ordinate corresponding to the penetration depth and the value of attenuation coefficient  $\alpha_0$ . This process is described by the formula:

$$F(t_i) = \sum_{k=-1}^G FA_k \exp(\alpha_0 \cdot f_k \cdot v \cdot t_i) \cdot \exp(-j \cdot 2 \cdot \pi \cdot f_k \cdot t_i) \quad (8)$$

where  $k$  stands for the index of the spectral component,  $f_k$  denotes frequency,  $FA_k$  is a complex spectrum of backscattered signal,  $v$  denotes phase velocity of the longitudinal acoustic wave in the skin and  $\alpha_0$  is the attenuation coefficient,  $t_i = i \cdot \delta t$  stands for time, where  $\delta t$  is a time step given by the signal sampling rate. The summation is carried over the whole range of frequencies of backscattered signal (G). The real part of  $F(\cdot)$  is the desired backscattered signal compensated for attenuation. Fig. 3 shows RF line before the TGC and attenuation compensations

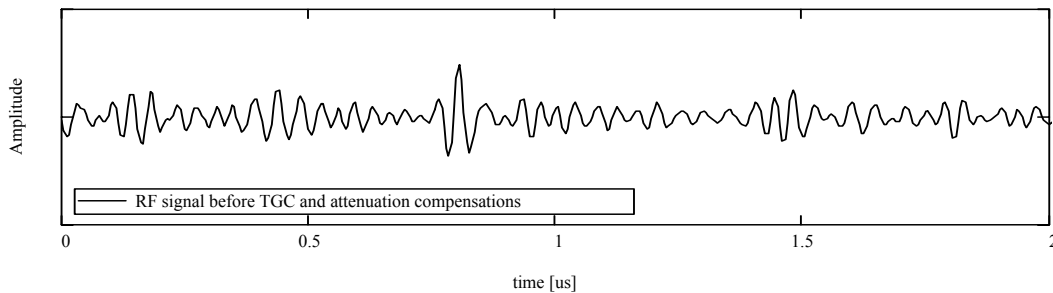


Fig.3. RF signal before the TGC and attenuation compensations.

After applying the compensation procedures the Hilbert transform was used to obtain signal envelope. Fig.4. shows signal after the TGC and attenuation compensation and its envelope.

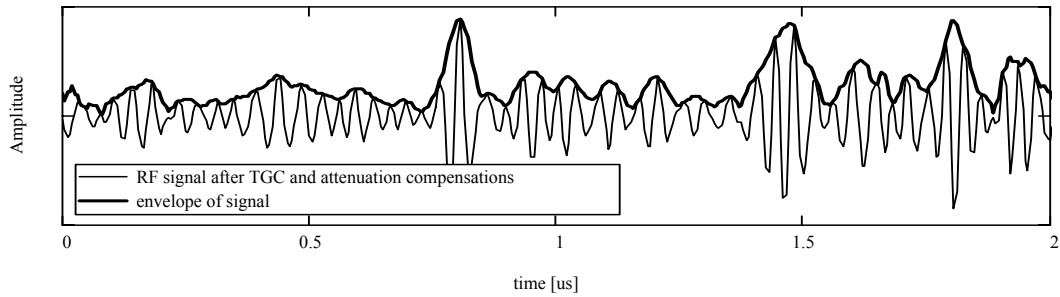


Fig.4. RF signal after the TGC and attenuation compensations and its envelope.

#### 4. RESULTS

This study presents the first results concerning the determination of the effective number of scatterers for the healthy human skin and skin with basal cell carcinoma.

To determine the applicability of the Rayleigh and K distributions in modeling the statistics of the ultrasonic backscatter from human dermis a comparison between experimental data and the K and Rayleigh distributions were done.

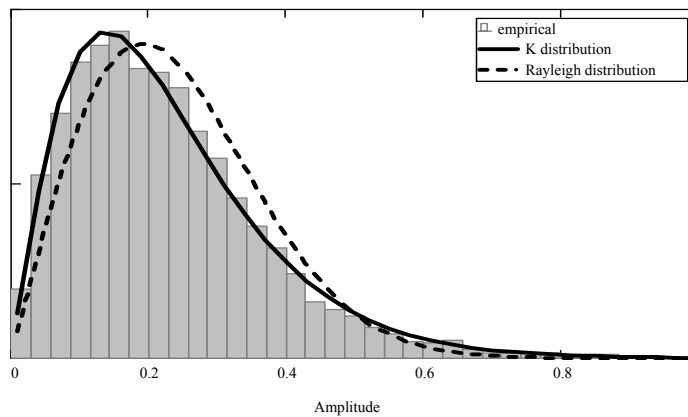


Fig.5. Theoretical Rayleigh and K distribution and histogram of empirical data obtained at the nape of neck.

Fig. 5 shows that the histogram of echo-envelope signal of the dermis is well described by K distribution, whereas the Rayleigh distribution does not successfully fit the experimental data. Table 1 presents the mean square error (MSE) values for the comparison empirical data and the K and Rayleigh distributions. On average, the error (mean MSE) between the K-distribution and empirical data was eight times lower than the error between the Rayleigh distribution and experimental data, pointing to a closer fit of data to the K-distribution.

Tab.1. MSE coefficient calculated for the histograms of the signal envelopes measured from the skin backscatter and theoretical K and Rayleigh (R) distribution.

	Skin in various parts of body										MEAN
	I case	II case	III case	IV case	V case	VI case	VII case	VIII case	IX case	X case	
K	0.014	0.011	0.013	0.009	0.011	0.013	0.011	0.008	0.014	0.021	<b>0.012</b>
R	0.087	0.085	0.09	0.084	0.076	0.081	0.087	0.089	0.073	0.078	<b>0.083</b>

Before the comparison could be undertaken, the characteristic parameter  $M$ , of the  $K$  distribution was calculated. The  $M$  parameter was calculated for every patient from region of skin with BCC and for comparison for region with non pathological changes of skin. The results obtained for four cases are presented in the table 2.

Tab.2.  $M$  coefficients calculated for experimental data obtained in vivo from the skin.

BCC	M	Healthy skin	M
I	1.117	I	1.989
II	1.165	II	1.611
III	0.908	III	1.611
IV	1.192	IV	1.929

The mean value of effective number of scatterers for healthy skin was equal 1.783 and ranged from 1.611 to 1.989. Pathological modifications of the tissue structure and cells resulting from the tumor modification that alters the tissue-ultrasound interaction can be potentially detected by evaluating the effective number of scatterers.

Fig.6a shows B-mode image of basal cell carcinoma, the white frame indicates the tumor masses. For comparison Fig. 6b presents healthy fragment of the skin of the same patient.

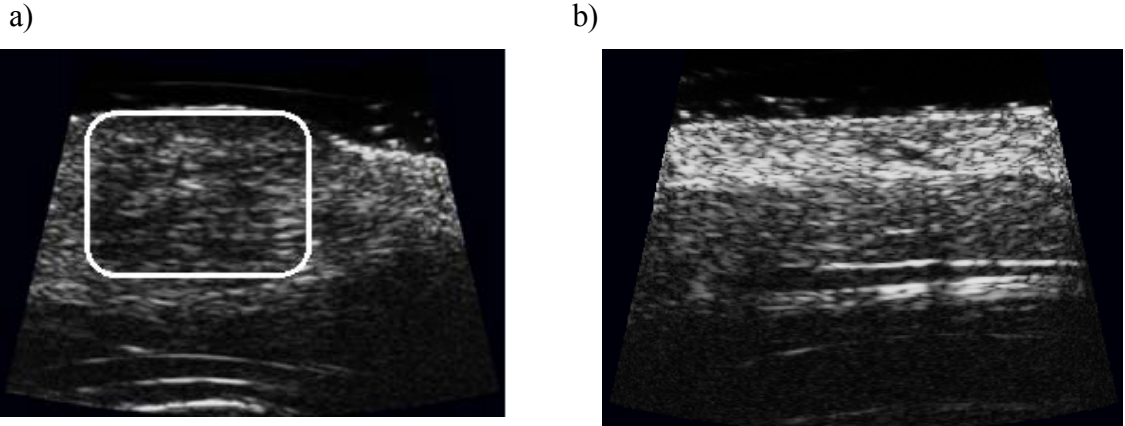


Fig.6. The B-mode image of basal cell carcinoma with marked tumor masses (a) and B-mode image of healthy human skin (b).

Fig.6 shows histograms determined for the above examples. It's easy to see that the shape of histograms is different. The values of analyzed parameters also were different for healthy skin and for skin with the BCC.

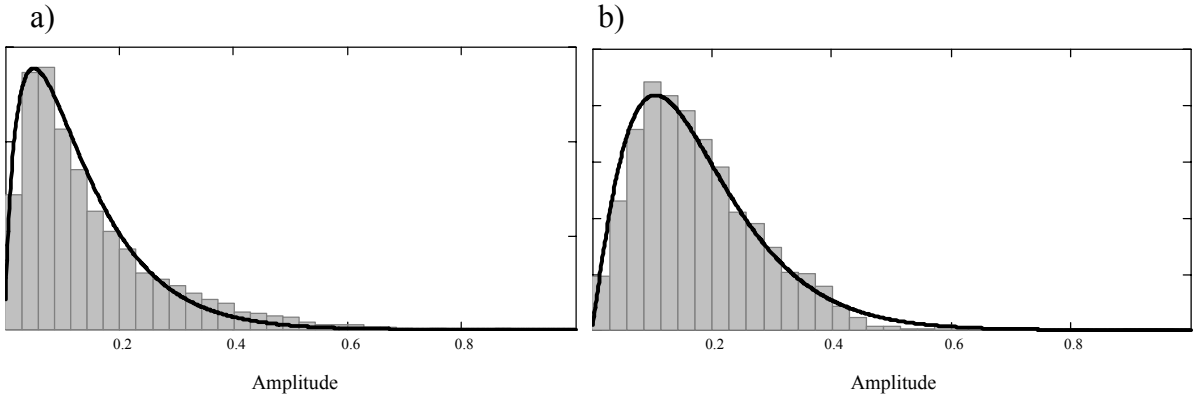


Fig.7. The histograms determined for skin with diagnosed Basal cell carcinoma (a) and for healthy skin (b).



This tendency was observed in all cases. For all patients values of effective number of scatterers obtained for skin lesions were lower than the values of M parameters calculated for healthy skin. The mean value of parameter M for skin lesions was equal 1.095 and ranged from 0.908 to 1.192.

## 5. CONCLUSION

The presented results are interesting and promising. We have found that the RF B-scans of skin obtained with our high frequency scanner can be used to characterize the tissue by evaluating its statistical parameters. The K-distribution, which is more dynamic model (better adjust to a wide range of scattering conditions than the narrowly-defined conditions that must be met for the Rayleigh model to be representative.), seems to be a good solution to describe the envelope statistics of signals from the human dermis. This more general model can provide additional useful information. Knowing that the envelope is closely related to the K distribution, the parameter M, associated with the effective number of scatterers, may provide meaningful information.

In our study M parameter obtained for tumor masses differs from those obtained for the healthy skin and it seems that in the future effective number of scatterers can be apply to the assessment of the skin structural properties.

## REFERENCES

- [1] V. Dutt, Statistical analysis of ultrasound envelope, The Mayo Graduate School, 1995.
- [2] J. Goodman, Speckle phenomena in optics theory and application, Roberts and Company Publishers, 2007.
- [3] E. Jakeman, P. Pusey, A model for Non-Rayleigh sea echo, IEEE Transactions on antennas and propagation, Vol. 24, 806-814, 1976.
- [4] E. Jakeman, R. Tough, Generalized K distribution: a statistical model for weak scattering, Journal Optics Society of America, Vol. 4, 1764-1772, 1987.
- [5] M. Lewandowski, A. Nowicki, High frequency coded imaging system with full software RF signal processing using Golay Transmission, IEEE International Ultrasonics Symposium, Vancouver, Canada, 2006.
- [6] J. Litniewski, Wykorzystanie fal ultradźwiękowych do oceny zmian struktury kości gąbczastej, Institute of Fundamental Technological Research of Polish Academy of Scienc 2006.
- [7] H. Piotrkowska *et al.*, Use of quantitative ultrasound to measure acoustic properties of human skin, Archives of Acoustics, Vol. 34, 471-480, 2009.
- [8] L. Weng *et al.*, Ultrasound speckle analysis based on the K distribution, Journal Acoustic Society of America, Vol.6, 2992-2995, 1991.
- [9] <http://www.mdblogger.com/2011/02/histology-and-explanation-of-basal-cell.html>.