ENDOBRONCHIAL TUMOR MASS INDICATION IN VIDEOBRONCHOSCOPY
Block based Analysis

Artur Przelaskowski, Rafal Jozwiak

Institute of Radioelectronics, Warsaw University of Technology, Nowowiejska 15/19, 00-665, Warsaw, Poland

Tomasz Zielinski

Department of Telecommunications, AGH University of Science and Technology, Al. Mickiewicza 30, 30-059, Krakow, Poland

Mariusz Duplaga

Collegium Medicum, Jagiellonian University in Krakow, Sw. Anny 12, 31-008, Krakow, Poland

Keywords: Bronchoscopy, Pattern recognition, Feature selection, Multiscale image processing.

Abstract: Computer-assisted interpretation of bronchial neoplastic lesion is an innovative but exceptionally challenging task due to highly diversified pathology appearance, video quality limitations and the role of subjective assessment of the endobronchial images. This work is focused on various manifestations of endobronchial tumors in acquired image sequences, bronchoscope navigation, artifacts, lightening and reflections, changing color dominants and unstable focus conditions. Proposed method of neoplasmatic areas indication was based on three steps of video analysis: a) informative frame selection, b) block-based unsupervised determining of enlarged textual activity, c) recognition of potentially tumor tissue, based on feature selection in different domains of transformed image and Support Vector Machine (SVM) classification. Prior to all of these procedures, wavelet-based image processing was applied to extract texture image for further analysis. Proposed method was verified with a reference image dataset containing diversified endobronchial tumor patterns. Obtained results reveal high accuracy for independent classification of individual (single video record) forms of endobronchial tumor patterns. The overall accuracy for whole dataset of 888 test blocks reached 100%. Less complex (approximately two times) procedure including initial blocks of interests selection reached accuracy of 96%.

1 INTRODUCTION

Recent advances in video technology enable for highly effective and safe diagnostic and therapeutic procedures of limited invasiveness, which is one of a key postulates in modern medicine. The quest for more sophisticated techniques in endoscopy is closely related to this trend (Duplaga, 2007). Even though, endoscopic examination remains stressful procedure for a patient and its outcome depends strongly on physician's skills and his subjective assessment of endoscopic images. Video recordings of endoscopic procedures stored in digital libraries are the source of diversified and often ambiguous information in terms of computer-assisted analysis. For example, video recordings of endoscopic examinations (e.g. gastroscopy, colonoscopy, bronchoscopy, etc.) contain not only images of normal and pathological endoluminal structures or diagnostic and therapeutic procedures, but also many poor quality or completely non–informative frames (e.g. blurred, out-of-focus, distorted, etc.).

1.1 Bronchoscopy

Bronchofiberoscopy is one of key diagnostic procedures employed in respiratory medicine and enabling direct visualization of the endoluminal structure of tracheobronchial tree. There are many indications when bronchofiberoscopy is performed, but in the

case of lung cancer suspicion it is obligatory. The procedure is usually accompanied by other diagnostic modalities enabling tissue sampling for pathologic evaluation. The progress in video technology had also its impact on bronchoscopy both in terms of available equipment and the scope of possible diagnostic and therapeutic techniques. Modern video bronchofiberoscopes contain an integral video camera system at the distal end and illumination system based on optical fibers assuring appropriate visibility of endobronchial structures. Endobronchial image is captured by the camera and displayed on the screen which can be conveniently placed in the bronchoscopy lab. The experience of performing physician considerably influences the effectiveness of the procedure and usually progress in the skill depends on the intensity of supervised training. The assessment of bronchoscopic images is poorly standardized and currently relies solely on procedure logbooks and subjective letters of competency (Bowling, 2007).

Video recording of bronchoscopic examinations demonstrate many common features with natural video sequences, e.g. general image features and natural content perception, color space, textural features, data dynamics, dominant objects properties. Different parts of bronchoscopic examinations differ in movement characteristics - from slow motion to dynamic video with fast camera movement across variable diagnostic content. Detection of pathological changes in bronchoscopy frame which comes after a sequence of normal images may be diagnostically challenging task.

Video recordings of bronchoscopy procedures are characterized by a considerable number of illegible sequences. Furthermore, great part of recorded frames brings images of normal tracheobronchial tree which must be also inspected during examination. While the content of the frames cannot be recognized and interpreted by any means because of limited acquisition conditioning or case-dependent specific image appearance, it is impossible to extract from it any diagnostic information. Such frames with unrecognized content were described as non-informative frames (Hwang, 2007). The appearance of these frames is highly diversified due to many factors influencing the quality of the endobronchial image. Essentially, we can distinguish: out-of-focus frames (occurred due to wrong camera position - too far/too close focus into/from mucosa of bronchi), blurred (motion blur due to rapid endoluminal camera movement), with sanguination (due to pathology presence or as a result of sampling of suspected tissues for pathology evaluation) and bubbled (as a result of camera lens cleaning). Informative frames carry poten-

tially important amount of diagnostic information.

However informative frames are not equal in the sense of visual quality, which change rapidly depending on current camera situation, position and movement. For example, any camera movement introduces versatile amount of motion blur (rapid or very fast camera movement can even lead to total loss of frame readability), while coverage of the camera lens with fluids or secretions results in loss of global focus condition. Despite of the diversified visual quality, most of informative frames collected in medical video library represent statistically normal images of different anatomical parts of tracheobronchial tree. The most significant video sequences in terms of pathological findings make relatively small part of registered frames. The appearance of various pathologic manifestations remains strongly highly diversified both inter-class (between different types of lesions) as well as intra-class (between different manifestations of the same type of pathology). Additionally every informative frame can be potentially affected by additional artifacts or distortions (e.g. specular reflections, lightening, etc.) which additionally hamper process of analysis. Examples of the contents of bronchoscopy video recordings, including non-informative and informative frames, different lesions manifestations and their diversification as well as possible distortions are presented in Fig. 1.

1.2 Pathology Recognition in Bronchoscopy Video

Computer-based tools developed for bronchoscopy support were designed with the objective of advanced visualization aspects like virtual bronchoscopy (Chung, 2006; Duplaga, 2005) or video (camera motion) tracking (Rai, 2006; Mori, 2002). Our work is focused on computer assisted automatic detection of lesions in bronchoscopy video. To the best of our knowledge, similar researches have been reported so far only for endoscopic modalities within gastrointestinal tract e.g. colonoscopy (Iakovidis, 2006), but not for bronchoscopy.

Among different pathologic endobronchial manifestations, detection of the features of the neoplastic process remains top interest. Endobronchial tumor mass constitutes relatively common manifestation of tumors affecting proximal parts of tracheobronchial tree - Fig. 2. Endobronchial tumor represents itself usually as a mass of different consistency in most cases with dense vasculature and having pink to purplish color.

The purpose of our work was to develop an efficient method for computer assisted interpretation

Figure 1: The examples of the content of bronchoscopy video. The first row shows examples of non-informative frames. The second row illustrates informative frames, readable with good visual quality. Example of onerous visual quality diversification is presented in the third row (adjacent frames from the same part of video examination). Finally, various bronchial lesions are presented in the fourth row (from left: tracheal stenosis, extravasations, widened main carina and mucous ulceration).

Figure 2: Examples of different endobronchial tumors. Pathology were outlined manually by experienced pulmonary medicine specialist.

of neoplastic lesions in bronchoscopy images. Suggested method exploits different image preprocessing techniques and is concentrated on texture-based image analysis. We proposed normalization of analyzed images, texture feature extraction in different domains (image, frequency, wavelet and contourlet), feature selection and unsupervised as well as supervised (SVM) classification, carried out at different stages of video content analysis. Implemented algorithms were optimized and verified experimentally.

2 MATERIAL AND METHODS

General task of diagnostically significant pattern recognition in bronchoscopic images is tough, challenging and technical conditions dependent because of mentioned above, seriously limited quality of image information. To succeed, recognition of endobronchial tumor mass was based on an analysis of bronchoscopy video frames according to precise successive selection of specified regions of interests (ROIs): informative frames, blocks of important content and blocks of potentially mass appearance.

Because dominant image feature used for recognition was texture, ROIs selection was based generally on textural data characteristics complimented with the factors of energy distribution in different data domains, entropy-based characteristics of stochastic information and other statistical measures of dominant signal trends.

2.1 Block-based Method Description

Proposed method of neoplasmatic areas indication is based on three steps of video analysis:

- IFS procedure, i.e. Informative Frame Selection, based on a whole frame characteristics of possible artifacts, unusual data and dominant texture recognition (SVM classifier);
- BUD method, i.e. Block-based Unsupervised Determination of enlarged textual activity areas in the IFS frames, based on energy distribution analysis and directional image characteristics in polar 2D Fourier space;
- BRT method, i.e. Block-based Recognition of Tumor tissue, based on feature extraction and 48 feature selection in image, frequency, wavelet, and contourlet domains, followed by the SVM classification.

Prior all of these successive procedures, wavelet-based image preprocessing was applied to extract texture image for further analysis.

2.1.1 Texture Extraction Preprocessing

A sequence of the following procedures was used:

- bronchoscopic frame normalization by successive conversions: a crop of source image area with mirror fulfilling of irregular corners, and conversion to grayscale image (G_I) - see Fig. 3;
- contrast-limited adaptive histogram equalization (CLAHE) of G_I with 4×4 pixel blocks and contrast function adjusted to 0.02: $CL_{4 \times 4, 0.02}(G\_I) = I_{AC}$;

- distorted wavelet synthesis of smoothed source image $I_S = \mathcal{W}^{-1}(\mathcal{W}(I_{AC}))$ and differential, texture image estimation as: $I_T = |I_{AC} - I_S|$, where $\mathcal{W}$ is 2-scale dyadic wavelet transformation based on non-perfect reconstruction filter bank;

- adaptive histogram equalization of texture image as follows: $I_{AT} = CL_{4\times4,0.02}(I_T)$.

Exemplary results of texture extraction procedure were presented in Fig. 3.
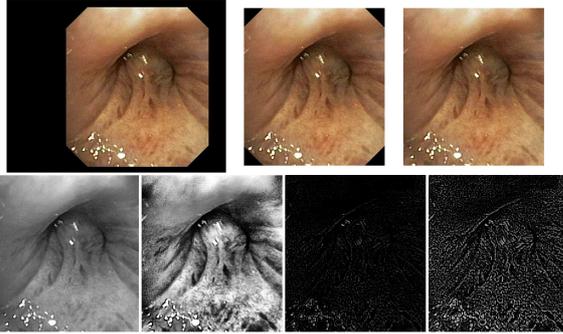
Figure 3: The effects of successive image preprocessing procedures: (left-right, top-down) source broncho frame, cropped image window, image with fulfilled corners, greyscale normalized version (G_I), adaptively enhanced G_I, textures extracted by wavelet-based procedure, texture enhanced by CLAHE.

Histogram equalization is used for imaging conditioning-invariance and texture extraction even in weak signal areas. CLAHE enhances local contrast of images by transforming the values into the intensity image. It operates on small data blocks so that the histogram of the output region approximately matches the uniform histogram. The neighboring blocks are bilineary interpolated to eliminate induced discontinuities on the block boundaries. The contrast, especially in homogeneous areas, is limited by contrast enhancement limit parameter to avoid amplifying the actual image noise or insignificant textures.

Multiresolution Signal Analysis according to wavelet-based concept is typically implemented with specific types of digital filter banks (FBs) known as two- channel perfect reconstruction (PR) filter banks. Those filters are associated with scaling functions (low pass one h) and the wavelets (high pass one g) of the transform kernel according to the following two equations (scaling and wavelet, respectively): $\phi(t) = \sqrt{2}\sum_n h_n \phi(2t-n)$ and $\psi(t) = \sqrt{2}\sum_n g_n \phi(2t-n)$. Conditions of the perfect reconstructions (Y is almost, i.e. according to the assumed precision, equal to X) with l delays for two-channel FB are as follows:

$$h(z)\tilde{h}(-z) + g(z)\tilde{g}(-z) = 0 \qquad (1a)$$

$$h(z)\tilde{h}(z) + g(z)\tilde{g}(z) = 2z^{-l} \qquad (1b)$$

Generally, wavelet decomposition requires the filters to be FIR (finite impulse response) and linearly phased to form orthogonal FBs. However, only Haar filter fulfill such requirements. Often used solution is biorthogonal FBs with insignificant redundancy of wavelet representation. For texture-oriented image processing we decided to design orthogonal FB by softening PR condition. The first term (eq. 1a) traditionally called the alias (cancellation) term is often fulfilled by using quadrature mirror filters (QMFs) with conditions: $h(z) = \tilde{g}(-z)$ and $g(z) = -\tilde{h}(-z)$, as we did. However, the second term (eq. 1b) called the distortion elimination term was used to control the distortion introduced in data processing to extract basic (lower frequency) signal content. Resulting filter proposition was spline non–PR FB with $h = [1/4, 1/2, 1/4]$ and miror $g = [-1/4, 1/2, -1/4]$ for signal smoothing by wavelet preprocessing (Przelaskowski, 2007).

2.1.2 Texture Characteristics and Recognition

We considered energy distribution characteristics across scales and subbands of wavelet domain diversified significantly classified tissue. Different classes of wavelet energy based features and histogram-based features from normalized wavelet coefficients were used. Moreover, entropy features (based on memoryless and joint source) for subbands compositions and homogeneity, correlation, energy and contrast of successive scale co-occurrence matrix of quantized coefficients were applied. SVM with optimized kernels and quality criteria was applied for classification and feature reduction procedures.

More precisely, the following textural features were estimated:

- in image domain:
  - statistical (variance, kurtosis, skewness, 0-order entropy, energy)
  - based on co-occurrence matrix (joint entropy, contrast, correlation, energy, homogeneity)
  - Tamura textural features (coarseness, directionality, contrast) (Tamura, 1978)

- in wavelet domain (symlet2 from nearly symmetrical wavelets, 2 scales of decomposition):
  - energy of approximation related to the energy of details
  - distribution of detail energy and entropy across scales

- – joint entropy of distribution if max magnitude details for successive scales
- – joint entropy, contrast, correlation, energy, homogeneity for co-occurrence matrix of quantized max magnitude details
- – max detail value of different subbands related to mean approximation energy
- in polar 2D Fourier domain:
  - – statistical moments of angle histogram of coefficient magnitudes
  - – energy of angle histogram
  - – parameters of polynomial approximation of angle histogram
- in contourlet domain – entropy of directional spectrum of two scales

Selected textural feature spaces were used for IFS and BRT procedures.

Informative frame selection (IFS) was based on a whole image I_{AT} characteristics for possible artifacts, regions of unusual data (non–informative) and informative regions with dominant texture recognition. 48 textural feature space was initially used for frame classification. Informative frames were extracted according to supervised classifier of over 800 training frames. SVM classifier with regularization and radial kernel was optimized for universal broncho applications.

Block-based unsupervised determination (BUD) of enlarged textual activity areas was used for fast defining of block of interests. Idea of second step of tumor extraction was high risk area segmentation through block-oriented cover of the image. Blocks of 50×50 pixels were verified basing on energy distribution analysis and directional image characteristics in polar 2D Fourier space. Two phase threshold selection: for energy distribution factors and directional factors was optimized for over 1000 test cases.

Block-based recognition of tumor tissue (BRT) was designed for final recognition of selected, active image blocks as potentially covered by tumor mass tissue. Recognition scheme was optimized for large test set of several thousand of test BUD blocks basing on textual feature extraction and selection for effective case classification. SVM procedure with regularization and radial kernel was used to classify each BUD block as diagnostically suspected of having pathology symptoms or normal.

Graphical form of selection and recognition results was used for indication of frame status as non–informative (purple mark in top left corner) or informative (lack of purple mark), while informative frame

regions could be covered by red (i.e. BUD) or yellow (i.e. BRT) blocks.

The proposed block-based method assisting intra-bronchial tumor mass recognition was implemented in MATLAB environment. For classification purpose we used standard SVM classifier from Matlab Bioinformatics Toolbox. Additionally we used some extra procedures form Contourlet Toolbox (Do, 2005), Beamlab (Donoho, 2001) and Wavelab (Buckheit, 1995).

2.2 Test Bronchoscopic Material

The proposed method was verified with a reference image data set containing different bronchoscopy video content exemplification (e.g. varied pathologies, diversified anatomical bronchi structures, therapeutic and diagnostic interventions, possible artifacts, etc.) clinically selected by experienced pulmonary medicine specialist from near 600 recorded and stored videos of bronchoscopy examinations. A reference set of close to 1300 images containing informative/non-informative frames was used to asses IFS efficiency.

For the purpose of BUD and BRT assessment 14 diversified cases of intrabronchial tumor manifestations were considered, each containing several especially selected frames with additional manual pathology outline. Every frame was divided into blocks of size 50x50, where each block was categorized to norm or pathology according to available physician outline. Near 900 selected blocks were finally used as a test set. Among these blocks, a small set of 38 blocks of one distinct case of intrabronchial tumor was selected. 16 of these blocks contained representative patterns of clear, high quality manifestation of the tumor completed with 22 blocks of surrounding, normal tissues. Such test sets were used for BUD/BRT verification and optimization.

3 EXPERIMENTS AND RESULTS

Examples of results achieved for the Matlab implementation of the described algorithm were presented and explained in Fig. 4. Illustration of the method adjusting to concrete examination and user requirements were outlined in Fig. 5.

IFS step tested against 1228 test frames gave even 100% accuracy of informative/non-informative selection with optimally adjusted radial kernel of SVM for complete space of extracted features. Optimization procedure indicated that IFS efficiency moderately
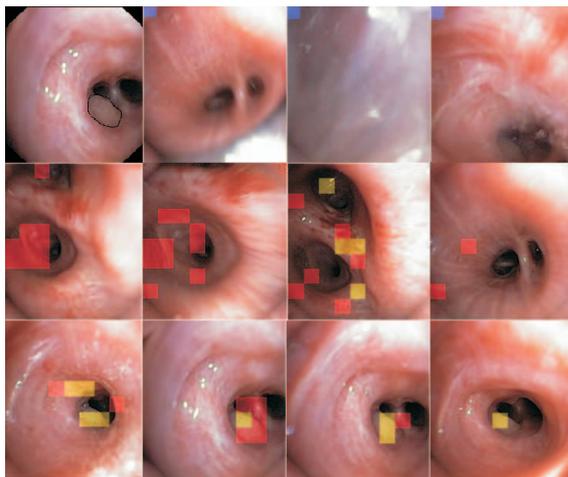
Figure 4: Exemplary results of the proposed block-based method applied to selected video examination with endobronchial tumor mass (reference frame with outlined pathology is presented in top left corner). Other frames of first row, marked with purple block, were classified as non-informative (first stage of algorithm). Red blocks represent areas with high texture information activity (related directly to second stage of proposed method) while yellow blocks marked out presence of pathology area (related directly to third stage of proposed method). Second row illustrates informative frames without clearly visible pathology (frames preceding appearance of pathology). Yellow blocks noticeable on third frame in this row (from left) represent false positive (FP) indications. Frames with visible pathology are present in third row. As we can see, pathology existence are indicated more or less precisely, but mainly correctly (yellow block are present in places corresponding to area outlined originally on reference frame). Some miss-classification examples can be seen on first frame (form left) - two yellow blocks located above pathology area.

Table 1: IFS effectiveness for selected sets of textural features and adjusted classifier; six the most useful features are: entropy of contourlet coefficients, ratio of approximation to detail wavelet coefficients, Tamura coarseness, entropy of wavelet details in successive scales, entropy and variance of extracted textures in image domain.

Number of the features	Classifier	Sensitivity	Accuracy
48	SVM/rbf	1	1
48	SVM/linear	.9	.92
24	SVM/rbf	1	1
24	SVM/linear	.89	.92
6	SVM/rbf	.91	.94
6	SVM/linear	0.87	0.87

depends on classification procedure and even significant reduction of the number of textural features. Exemplar results were presented in Tab. 1.

Sensitivity of BUD method was optimized in relation to high enough specificity through adjusting of
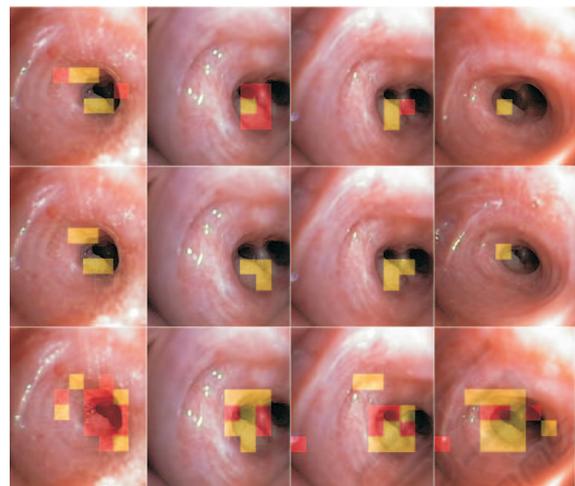
Figure 5: Impact of method parametrization for overall detection efficiency. Each row illustrates results for different set of method parameters. Depending on the selected set of initial presets proposed method is more or less sensitive. Selection of this parameters should set a compromise between method sensitivity and specificity.

Table 2: Adjusting of BUD parameters to select the best balance between sensitivity and specificity of the method. Threshold values of 4.35/1 (4.35 for normalized energy distribution and 1 for directional image characteristics) were used finally.

Two threshold values of BUD	Sensitivity	Specificity
4.35/4	.65	.58
4/4	.70	.36
4.35/3	.81	.44
4.35/1	.89	.38
4/2	.95	.1

two selective thresholds. Diversified in local energy test set of 888 image blocks was used to make the procedure more universal, sufficiently efficient as on-line detector of suspicious to tumor areas and selective enough for more precise supervising of BRT procedure.

Satisfying sensitivity of 94% with specificity up to 86% was achieved for more distinct test set of 38 selected blocks. The balance range between sensitivity and specificity for total test set of 888 blocks was presented in Tab. 2. Maximum accuracy was limited to 61%.

Accuracy of BRT method was adjusted to 98% for 888 test control or pathology blocks. Selected 30 textural feature space was used. Six the most useful features are: entropy of contourlet coefficients, mean of angle histogram and the parameter of polynomial approximation of angle histogram in polar 2D Fourier domain, mean local entropy in image-texture domain

combined with texture energy, global entropy and energy in image-texture domain. However, accuracy for these 6 features was only 75%. Accuracy of BRT for more flexible and universal linear kernel of SVM was only 76%.

Obtained results reveal high accuracy for independent classification of individual differential forms of endobronchial tumor patterns, especially basing on time consuming IFS-BRT procedure. The overall accuracy for whole dataset of 888 test blocks reached 100%. Thus, time consuming IFS-BRT combination that assumed feature extraction and SVM classification for each frame and next successively for all frame blocks is effective enough to fit classification rules to tumor detection problem. Less complex (approximately five times) procedure of complete IFS-BUD-BRT reached accuracy of 96%.

All verified procedures were designed to analyze bronchoscopic video in order to indicate the blocks of high susceptibility to tumor mass. A way of frame and frame block selection to be analyzed depends on application requirements. Because of computational complexity of BRT, which is fundamental procedure for tumor recognition, ad-hoc section method of blocks of interests, similar to BUD or other interactive methods based on human-computer interfaces, are useful for close to on-line application.

4 CONCLUSIONS

Clinical usefulness of the proposed method should be further tested in conditions of bronchoscopy suit. Reliable experimental procedure strongly depends on significantly diversified technical conditions of bronchofiberoscopes and test cases. Moreover, the feasibility of this method may be affected by the limited standardization of the procedure and significant role of subjective assessment. However, automatic indications in almost on-line mode (second stage of the method) or more reliable in off-line mode (full method application) seems to be useful as an assistant tool for more careful bronchoscopic video analysis. Objectified indicators of special regions of interests are useful for standardized protocol design, comparative analysis and education of inexperienced doctors. As far as the authors are aware, it is the first attempt of development of the tool based on the automatic probable pathology indication supporting bronchoscopic examination.

REFERENCES

Duplaga, M., Leszczuk, M., Przelaskowski, A., Janowski, L. and Zieliski, T. (2007). Bronchovid - zintegrowany system wspomagajcy diagnostyk bronchoskopow. *Przegld Lekarski 64:42-48.*

Bowling, M., Downie, G., Wahidi, M. and Conforti, J. (2007). Self-Assessment Of Bronchoscopic Skills In First Year Pulmonary Fellows. *Chest Vol. 132, Issue 4.*

Hwang, S., Oh, J., Lee, J., Tavanapong, W., de Groen, P. C. and Wong, J. (2007). Informative Frame Classification for Endoscopy Video. *Medical Image Analysis Vol. 11, No 2:100-127.*

Chung, A. J., Deligianni, F., Shah, P., Wells, A. and Yang, G. Z. (2006). Patient Specific Bronchoscopy Visualisation through BRDF Estimation and Disocclusion Correction. *IEEE Transactions of Medical Imaging 25(4):503- 513.*

Duplaga, M. and Socha, M. (2005). Aplikacja oparta na bibliotece VTK wspomagajca zabiegi bronchoskopowe. *Bio-Algorithms and Med-Systems I(l/2):191-196.*

Rai, L., Merritt, S. A. and Higgins, W. E. (2006). Real-time image-based guidance method for lung-cancer assessment. *IEEE Conf. Computer Vision and Pattern Recognition 2:2437-2444.*

Mori, K., Deguchi, D., Sugiyama, J., Suenaga, Y., Toriwaki, J., Maurer, C. R. Jr, Takabatake, H. and Natori, H. (2005). Tracking of a bronchoscope using epipolar geometry analysis and intensity-based image registration of real and virtual endoscopic images. *Med. Image Anal. 6:321-365.*

Iakovidis, D. K., Maroulis, D. E. and Karkanis, S. A. (2006). An Intelligent System for Automatic Detection of Gastrointestinal Adenomas in Video Endoscopy *Computers in Biology and Medicine. Vol. 36, 10:1084-1103.*

Przelaskowski, A., Bargiel, P., Sklinda K. and Zwierzynska E. (2007). Ischemic stroke modeling: multiscale extraction of hypodense signs *Lecture Notes in Artificial Intelligence 4482:171-181, Springer Verlag.*

Tamura, H., Mori, S. and Yamawaki, T. (1978). Textural features corresponding to visual perception *IEEE Trans. Systems, Man. and Cybern. Vol. 8, 6:460-472 .*

Do, M. N. and Vetterli, M. (2005). The contourlet transform: an efficient directional multiresolution image representation *IEEE Trans Image Proces. Vol. 14, 12:2091-2106 .*

Donoho, D. L. and Huo, X. (2001). Beamlets and Multiscale Image Analysis *Computational Science and Engineering, Multiscale and Multiresolution Methods, Springer.*

Buckheit, J. B. and Donoho, D. L. (2005). WaveLab and Reproducible Research *Dept. of Statistics, Stanford University, Tech. Rep. 474.*