Image Analysis System For Assessing The Estrogen Receptor's Positive Status In Breast Tissue Carcinomas

Spiros Kostopoulos, Dionisis Cavouras, Antonis Daskalakis, Panagiota Ravazoula

and George Nikiforidis

***Abstract*—Estrogen receptor (ER) status has proven to be a significant factor for the prediction of clinical response to hormonal therapy, in patients with breast cancer. In clinical practice, assessment of ER positive status relies on the subjective identification of the expressed nuclei in the specimens. The aim of this study was the development of a computer-aided image analysis system, employing an unsupervised segmentation algorithm based on the *L*a*b* color space transformation. Kendall's coefficient of concordance showed an adequate level of agreement (Kendall's W=0.79) between the clinical evaluation of the physician and the objective quantification of the automatic computer-aided system. Computer-assisted determination of ER status may be used as a second opinion tool in routine assessment of immunohistochemical sections.**

I. Introduction

BREAST carcinoma is the most common malignancy among females with an increasing tendency [1]-[3]. Estrogens act on breast tissue (normal and malignant), by binding to estrogen receptors (ER), and cause cell proliferation. Although this process is important for normal breast development, it includes an inherent risk of developing cancer cells. Patients with breast cancer cells that express ER in their nuclei (ER+ status) undergo different therapeutic management and treatment from those that breast cancer cells do not possess ER (ER- status) [4], [5]. It has been shown that the ER status is a significant biologic factor for the prediction of clinical response to hormonal therapy [6].

Recently, evaluation of ER status is performed by means of immunohistochemistry (IHC) [7], [8]. In clinical practice, the histopathologist chooses the IHC stained sections to be assessed, based on the diagnostic assessment of hematoxylin and eosin stained slides [4]. Assessment of ER positive status relies on the subjective identification of the percentage positive (stained) nuclei, under microscopic review. Due to the intrinsic inter- and intra-observer variability, an objective quantification is required for the accurate determination of the ER status [9].

Computer-aided image analysis systems have been proposed for the objective quantification of ER status as second opinion tools. Previous studies are mainly concerned with the employment of commercially developed image analysis systems [9]-[14] that mostly employ global thresholding techniques. Schnorrenberg *et al.* [15] have proposed a promising approach for the accurate detection of ER status, utilizing an in-build algorithm for the detection and classification of individual nuclei, however, requiring user interaction.

The present study is focused on the development of an automatic computer-aided image analysis system for the objective assessment of ER positive status of breast cancer with no user interaction. For this, a specific color transformation followed by unsupervised clustering were developed to separate positively stained (brown) from negatively stained (blue) nuclei and from background tissue and, thus, to automatically assess the percentage of positive nuclei (ER+ status) present in the IHC stained image. Results were compared against the physician's objective evaluation.

II. Material And Methods

Twenty nine immunohistochemically stained specimens of breast cancer were collected by an experienced histopathologist (P.R) from the Department of Pathology of the University Hospital of Patras, Greece. For each specimen, the ER expression was semi-quantitative assessed (P.R), based on a clinical scoring protocol [16]. According to this protocol, the percentage of the number of positive stained nuclei to the total number of positive and negative nuclei was visually inspected (Table I), from a representative region, where a large number of positive nuclei existed. Brown and blue nuclei were regarded as positive and negative stained respectively. Five percent was

TABLE I
THE SCORING SYSTEM USED IN THIS STUDY

Score	Proportion of positive nuclear staining
1	6-33%
2	34-66%
3	67-100%

Manuscript received June 30, 2006. This work was supported by the Greek State Scholarship Foundation (IKY).

S. Kostopoulos, A. Daskalakis and G. Nikiforidis are with the Medical Image Processing and Analysis Group, Department of Medical Physics, School of Medicine, University of Patras, 26500 Rio, Greece (2610-997745; e-mail: skostopoulos@upatras.gr).

D. Cavouras is with Medical Image ans Signal Processing Laboratory, Department of Medical Instruments Technology, Technological Educational Institute of Athens, 12210 Athens, Greece (e-mail: cavouras@teiath.gr).

P. Ravazoula is with the Department of Pathology, University Hospital of Patras, 26500 Rio, Greece.

used as the cutoff value of positive status. The studied cases had a physician's positivity score of 1 (11/28) and 2 (18/28).

For each case, a number of images (mean 7, range 6-9) were selected from a predefined region and were digitized (1300x1030x16bit) at a magnification of x400 using a light Zeiss Axiostar-Plus microscope (ZEISS; Germany) and a Leica DC 300F color video camera (LEICA; Germany). Each digitized image was stored in an uncompressed tagged image format file (TIFF).

The original colored (RGB) image was converted to L*a*b* (CIELAB) color space (Figure 1 a, b). L* represents the difference between light (100) and dark (0), whereas the other two coordinates a* and b* represent redness (-a*)-greenness (+a*) and blueness (-b*)-yellowness (+b*) respectively. Thus, the color information was confined in the 2d color space instead of the 3d (RGB) [17].

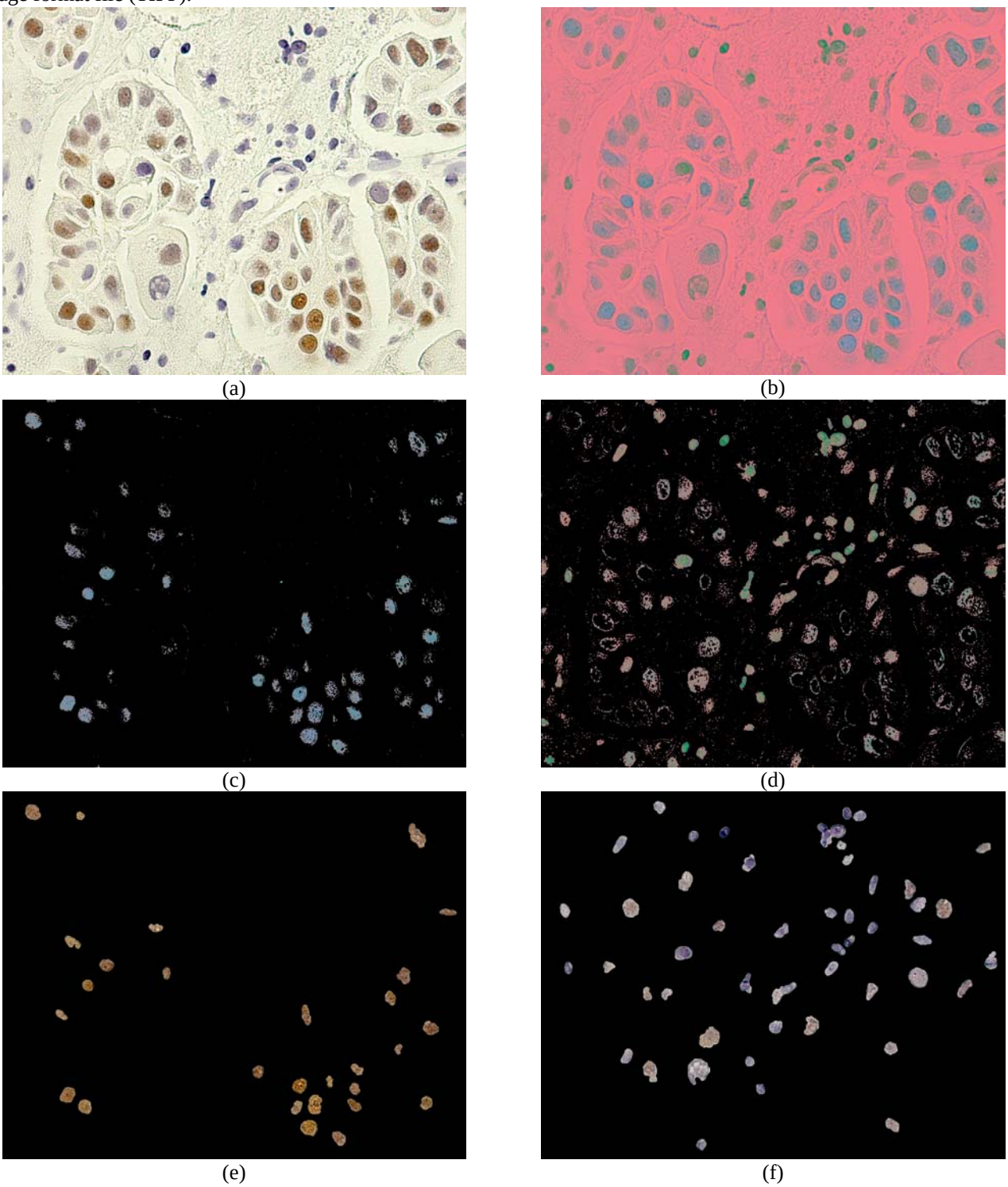

Fig.1. (a, b) Original image in RGB and L*a*b color space respectively. (c, d) brown and blue pixels respectively resulted from the k-means algorithm. (e, f) final segmented images with brown and blue objects/nuclei respectively.

The (a* and b*) image pixels were then fed to the k-means clustering algorithm (k=3). The algorithm [18] follows an iterative procedure, where its pixels are assigned to the cluster with the minimum Euclidean distance from the cluster's centroid. The aim was to divide the image pixels into three clusters (one for the 'brown' pixels, one for the 'blue' and a third for the background pixels) giving as output the centroids of the three clusters and the cluster label of each pixel. The 'blue' and 'brown' clusters were experimentally determined having the second and third largest centroid while the smallest centroid resided in the background cluster. Consequently, two images resulted, one having the brown and the other the blue nuclei (Figure 1 c, d). Images were further processed with fill holes, morphological open/closing, and size filters operations [19] to eliminate small, noisy regions and to omit corrupted nuclei across image boundaries (Figure 1, e, f).

From all captured images of the same specimen, the ER positive status was calculated as the percentage of all brown nuclei over the total number of nuclei present.

III. Results

A mean value of 430 nuclei with a range of 350-590 was segmented for each specimen. Table II shows the self-evaluation performance in identified brown and blue nuclei from one case. In 329 brown and 362 blue nuclei, the system was correctly identified 87.5% and 81.5% nuclei respectively.

TABLE II
SELF-EVALUATION

	Correctly identified nuclei	Lost nuclei	Accuracy
Brown nuclei	*288*	*41*	87.5%
Blue nuclei	*295*	*67*	81.5%

Self-evaluation of the computer-aided system in identifying brown and blue nuclei.

Table III shows the percentage ER positive status as resulted by the computer-aided system against the physician's estimation, as well the absolute differences and the corresponding score.

TABLE III
RESULTS OF THE PERCENTAGE ER POSITIVE STATUS EVALUATION AND THE CORRESPONDING SCORES

	Physician's evaluation (%)	Computer-aided system evaluation (%)	Absolute differences (%)	ER score Physician	ER score System
1	20	12	8	1	1
2	20	31	11	1	1
3	20	15	5	1	1
4	20	13	7	1	1
5	20	11	9	1	1
6	20	12	8	1	1
7	20	25	5	1	1
8	20	32	12	1	1
9	30	41	11	1	2
10	30	21	9	1	1
11	30	18	12	1	1
12	40	26	14	2	1
13	40	51	11	2	2
14	40	35	5	2	2
15	40	49	9	2	2
16	40	54	14	2	2
17	40	25	15	2	1
18	50	37	13	2	2
19	50	39	11	2	2
20	50	60	10	2	2
21	50	45	5	2	2
22	50	45	5	2	2
23	50	39	11	2	2
24	50	65	15	2	2
25	50	60	10	2	2
26	50	39	11	2	2
27	50	42	8	2	2
28	50	41	9	2	2
29	50	38	12	2	2

Wilcoxon rank test revealed that no significant differences existed between the two percentages evaluations (P=0.38). A mean value of 9.8% of the percentage absolute differences in Table III existed in examined cases.

Table IV presents the results of the computer-aided ER evaluation system against the physician's scores. Computer-aided image analysis system resulted in 89.6% overall accuracy, ranking correctly 26/29 cases. 90.9% (10/11) of

the specimens, having an ER positive score of 1, were correctly scored. Likewise, 88.9% (16/18) of those having a score of 2 were correctly evaluated.

TABLE IV
RESULTS OF THE COMPUTER-AIDED ER EVALUATION SYSTEM AGAINST PHYSICIAN'S SCORE

	Computer–aided score		
ER Score	**1**	**2**	**Accuracy**
1	*10*	*1*	90.9%
2	*2*	*16*	88.9%
Overall accuracy			89.6%

Kendall's coefficient of concordance (KCC) was used to determine the level of agreement between assessments [20-22]. KCC ranges from 0 to 1. The higher the value the stronger is the association among ratings. KCC revealed adequate level of agreement among the physician and the computer-aided system ($W=0.79$, $p<0.05$).

IV. DISCUSSION

An unsupervised algorithm for ER positive status assessment was introduced and tested against physician's estimation. Differences in percentage evaluation (Table III) may be due to lost nuclei by the proposed system (see Table II) and/or due to the physician's subjective assessment. Nuclei are mainly missed by the system as a result of morphological and size filters operations immediately after unsupervised clustering. It has to be noted that only cases of intermediate ER positive status were considered (Table IV), since such cases contain similarly distributed blue and brown nuclei, rendering the problem of ER status determination more complex.

A variety of commercially available image analysis systems have been studied for quantification of IHC sections. Recently, Mofidi et. al [10] have used a well known image analysis software and a similar scoring methodology with the presented study, reporting significant correlation of ER status between manually and image analysis assessment, even though, subjectivity was introduced in specifying color characteristics. Schnorrenberg et al [15] have utilized an in-house algorithm for the detection and classification of individual nuclei taking into account a scoring system that considered both nuclei intensity and percentage positivity. Their work revealed correlated results with physicians, though limited to user interaction.

Estimation of ER status in everyday clinical practice has been shown to be useful for its prognostic and therapeutic importance. Even so, the diagnostic consensus of a number of physicians is essential for superior accuracy [23]. In consequence, a computer-aided image analysis system, as the one proposed in the present study, may provide a useful opinion tool, in clinical assessment of IHC sections.

REFERENCES

[1] M. V. Karamouzis, E. Likaki-Karatza, P. Ravazoula, F. A. Barda, D. Koukouras, E. Tzorakoleftherakis, A. G. Papavassiliou, and H. P. Kalofonos, "Non-Palpable Breast Carcinomas: Correlation Of Mammographically Detected Malignant-Appearing Microcalcifications And Molecular Prognostic Factors," *Int J. Cancer*, vol. 102, pp. 86-90, 2002.

[2] B. W. Stewart and P. Kleihues, *World Cancer Report*. Lyon: IARC Press, 2003.

[3] S. Sunpaweravong and P. Sunpaweravong, "Recent Developments in Critical Genes in the Molecular Biology of Breast Cancer," *Asian Journal Of Surgery*, vol. 28, pp. 71-75, 2005.

[4] L. K. Diaz and N. Sneige, "Estrogen receptor analysis for breast cancer: current issues and keys to increasing testing accuracy," *Adv Anat Pathol*, vol. 12, pp. 10-9, 2005.

[5] Y. Kun, L. C. How, T. P. Hoon, V. B. Bajic, T. S. Lam, A. Aggarwal, H. G. Sze, W. S. Bok, W. C. Yin, and P. Tan, "Classifying the estrogen receptor status of breast cancers by expression profiles reveals a poor prognosis subpopulation exhibiting high expression of the ERBB2 receptor," *Hum Mol Genet*, vol. 12, pp. 3245-58, 2003.

[6] B. Jasani, A. Douglas-Jones, A. Rhodes, S. Wozniak, P. J. Barrett-Lee, J. Gee, and R. Nicholson, "Measurement of estrogen receptor status by immunocytochemistry in paraffin wax sections," *Methods Mol Med*, vol. 120, pp. 127-46, 2006.

[7] A. A. Thike, M. J. Chng, S. Fook-Chong, and P. H. Tan, "Immunohistochemical expression of hormone receptors in invasive breast carcinoma: correlation of results of H-score with pathological parameters," *Pathology*, vol. 33, pp. 21-5, 2001.

[8] J. M. Harvey, G. M. Clark, C. K. Osborne, and D. C. Allred, "Estrogen receptor status by immunohistochemistry is superior to the ligand-binding assay for predicting response to adjuvant endocrine therapy in breast cancer," *J Clin Oncol*, vol. 17, pp. 1474-81, 1999.

[9] L. K. Diaz, A. Sahin, and N. Sneige, "Interobserver agreement for estrogen receptor immunohistochemical analysis in breast cancer: a comparison of manual and computer-assisted scoring methods," *Ann Diagn Pathol*, vol. 8, pp. 23-7, 2004.

[10] R. Mofidi, R. Walsh, P. F. Ridgway, T. Crotty, E. W. McDermott, T. V. Keaveny, M. J. Duffy, A. D. Hill, and N. O'Higgins, "Objective measurement of breast cancer oestrogen receptor status through digital image analysis," *Eur J Surg Oncol*, vol. 29, pp. 20-4, 2003.

[11] H. A. Lehr, D. A. Mankoff, D. Corwin, G. Santeusanio, and A. M. Gown, "Application of photoshop-based image analysis to quantification of hormone receptor expression in breast cancer," *J Histochem Cytochem*, vol. 45, pp. 1559-65, 1997.

[12] S. V. Makkink-Nombrado, J. P. Baak, L. Schuurmans, J. W. Theeuwes, and T. van der Aa, "Quantitative immunohistochemistry using the CAS 200/486 image analysis system in invasive breast carcinoma: a reproducibility study," *Anal Cell Pathol*, vol. 8, pp. 227-45, 1995.

[13] S. Bacus, J. L. Flowers, M. F. Press, J. W. Bacus, and K. S. McCarty, Jr., "The evaluation of estrogen receptor in primary breast carcinoma by computer-assisted image analysis," *Am J Clin Pathol*, vol. 90, pp. 233-9, 1988.

[14] C. Charpin, P. M. Martin, B. De Victor, M. N. Lavaut, M. C. Habib, L. Andrac, and M. Toga, "Multiparametric study (SAMBA 200) of estrogen receptor immunocytochemical assay in 400 human breast carcinomas: analysis of estrogen receptor distribution heterogeneity in tissues and correlations with dextran coated charcoal assays and morphological data," *Cancer Res*, vol. 48, pp. 1578-86, 1988.

[15] F. Schnorrenberg, C. S. Pattichis, K. C. Kyriacou, and C. N. Schizas, "Computer-Aided Detection Of Breast Cancer Nuclei," *IEEE Transactions On Information Technology In Biomedicine*, vol. 1, 1997.

[16] R. Leake, D. Barnes, S. Pinder, I. Ellis, L. Anderson, T. Anderson, R. Adamson, T. Rhodes, K. Miller, and R. Walker, "Immunohistochemical detection of steroid receptors in breast cancer: a working protocol. UK Receptor Group, UK NEQAS, The Scottish Breast Cancer Pathology Group, and The Receptor and Biomarker Study Group of the EORTC," *J Clin Pathol*, vol. 53, pp. 634-5, 2000.

[17] L. Xu, M. Jackowski, A. Goshtasby, D. Roseman, S. Bines, C. Yu, A. Dhawan, and A. Huntley, "Segmentation of skin cancer images," *Image and Vision Computing*, vol. 17, pp. 65-74, 1999.
[18] A. K. Jain and R. C. Dubes, *Algorithms for Clustering Data*. New Jersey: Prentice-Hall, 1988.
[19] P. Spyridonos, D. Cavouras, P. Ravazoula, and G. Nikiforidis, "Neural network-based segmentation and classification system for automated grading of histologic sections of bladder carcinoma," *Anal Quant Cytol Histol*, vol. 24, pp. 317-24, 2002.
[20] J. L. Fleiss, *Statistical Methods for Rates and Proportions*, 2nd ed. NY: Wiley & Sons, 1981.
[21] A. Rhodes, B. Jasani, D. M. Barnes, L. G. Bobrow, and K. D. Miller, "Reliability of immunohistochemical demonstration of oestrogen receptors in routine practice: interlaboratory variance in the sensitivity of detection and evaluation of scoring systems," *J Clin Pathol*, vol. 53, pp. 125-30, 2000.
[22] M. Craigie, B. Loader, R. Burrows, and S. Muncer, "Reliability of health information on the Internet: an examination of experts' ratings," *J Med Internet Res*, vol. 4, pp. e2, 2002.
[23] R. Jagoe, J. H. Steel, V. Vucicevic, N. Alexander, S. Van Noorden, R. Wootton, and J. M. Polak, "Observer variation in quantification of immunocytochemistry by image analysis," *Histochem J*, vol. 23, pp. 541-7, 1991.