Georgia D. Tourassi, PhD

Journey toward Computer-aided Diagnosis: Role of Image Texture Analysis[1]

Currently, methods of image texture analysis are undergoing great development and utilization within the field of medical imaging. Given the general interest and striking growth in computer-aided diagnosis (CAD), the application of texture analysis in the diagnostic interpretation of radiologic images has become a rapidly expanding field of research.

The article by Chen et al (1) that appears in this issue of *Radiology* further demonstrates the potential of image texture analysis as an important component of CAD algorithms. Specifically, the authors present a CAD tool that was developed for the characterization of solid breast nodules as benign or malignant on ultrasonographic (US) images. The proposed diagnostic scheme follows a familiar two-step approach. Initially, textural information is extracted from the image. Subsequently, the extracted information is fed into a decisional algorithm (eg, an artificial neural network) that is designed to perform the diagnostic task. The clinical importance of the diagnostic problem and the high accuracy rate reported by the authors make their article interesting to those involved in the diagnosis of breast cancer.

However, from a methodological point of view, the main attraction of this study is that the combination of image texture analysis and automated decision making offers a promising approach to a clinical challenge (2). Successful applications of the above CAD strategy have been reported for other types of medical images (3–12). While further development and testing is required to establish the true clinical effect of such decision-support systems, texture analysis appears to open a new exciting path in the journey toward CAD in radiology. Consequently, two questions are raised: What is the realistic contribution of texture analysis in the computer-aided interpretation of medical images, and to what extent can it be expected to improve interpretative accuracy?

Index terms:
Breast neoplasms, diagnosis, 00.30
Breast neoplasms, US, 00.1298
Computers, diagnostic aid
Editorials
Images, analysis
Images, interpretation
Ultrasound (US), tissue characterization, 00.1298

Abbreviation:
CAD = computer-aided diagnosis

Radiology 1999; 213:317–320

[1] From the Department of Radiology, Duke University Medical Center, Box 3302, Erwin Rd, Durham, NC 27710. Received July 6, 1999; accepted July 16. **Address reprint requests to** the author (e-mail: *gt@deckard.mc.duke.edu*).

See also the article by Chen et al (pp 407–412) in this issue.

Deconstructing the Diagnostic Process

The diagnostic interpretation of medical images is a multifaceted task. Its objective is the accurate detection and precise characterization of potential abnormalities—a crucial step toward the institution of effective treatment. Achieving this goal relies on the radiologists' successful integration of two distinct processes: (*a*) the process of image perception to recognize unique image patterns and (*b*) the process of reasoning to identify relationships between perceived patterns and possible diagnoses. Both processes depend heavily on the radiologists' empirical knowledge, memory, intuition, and diligence. Unquestionably, the radiologists approach the diagnostic task with a level of intelligence, flexibility, and common sense that is difficult to duplicate with a computer. Thus far, research findings have demonstrated only that computers are the best savants when it comes to executing mathematically rigorous steps toward the solution of narrowly defined problems.

Nonetheless, the radiologist's approach is not devoid of limitations. There are well-documented errors and variations in the human interpretation of clinical images (13). Some study findings even indicate that the same errors are being made now, as they were in earlier decades (14–16). In addition, findings from a recent review of 20 years experience in malpractice litigation in radiology has shown that the overwhelming majority of legal cases were due to alleged diagnostic mistakes that were attributed to perceptual errors (17,18) and errors in judgment (19). In this context, computers can be our allies if we deconstruct the complex processes in image perception and diagnostic reasoning into a series of single, well-defined tasks—tasks for which computers have been proved to be extremely suited and with which they have been successful. But how can we do that?

Texture Analysis: More than Meets the Eye?

The impressive performance of humans in processing visual information is a constant source of inspiration for researchers who are trying to understand, to model, and even to duplicate visual learning. Despite decades of intensive research in biologic visual systems and perceptual intelligence, there is a limited understanding and an ongoing debate about the basic mechanisms that underlie visual perception. In contrast, there is universal agreement that texture is a rich source of visual information and is a key component in image analysis and understanding in humans (20). Texture is known to provide cues about scenic depth and surface orientation and, as such, describes the content of both natural and artificial images. Also, there is evidence of perceptual learning in texture-coding mechanisms (21,22) and in textural discrimination (23).

In light of this, engineers have focused their attention on developing algorithms that can quantify the textural properties of an image. (Even though there is no formal definition of the term, "optical texture" is used to describe the local spatial variations in image brightness, which, consequently, are related to properties such as coarseness and regularity.) Empirical evidence and researchers' inventiveness have led to numerous algorithms for texture analysis. The available algorithms typically differ in the type of image information that is captured and in the coding mechanism. For example, there are structural approaches that depict deterministic texture as a hierarchy of spatial arrangements of well-defined primitives. There are statistical approaches that represent texture with the nondeterministic properties of the relationships between the gray levels on images. Some techniques represent texture on the basis of the spectral properties of an image. Others are model-based techniques that analyze texture by identifying an appropriate model that reflects the prior beliefs and knowledge about the type of images to be analyzed. There are textural features that describe local image statistics and others that describe global statistics. Also, there are more sophisticated textural features determined with Markov random-field models that relate local image structures to specific global constraints.

The evolution and diversity of available techniques for texture analysis are a testament to the advancement of this field. An extensive survey of textural definitions, models, and analytic algorithms can be found elsewhere (24). Texture analysis is ultimately concerned with automated methods that can derive image information from a purely computational point of view. As such, it is nothing more than another type of numeric manipulation of digital or digitized images to get quantitative measurements. But contrary to the discrimination of morphologic information (ie, shape, size), there is evidence that the human visual system has difficulty in the discrimination of textural information that is related to higher-order statistics or spectral properties on an image (25,26). Consequently, texture analysis can potentially augment the visual skills of the radiologist by extracting image features that may be relevant to the diagnostic problem but that are not necessarily visually extractable.

The idea to use texture analysis in medical imaging has been considered since the early 1970s (27). However, the exciting evolution of both texture analysis algorithms and computer technology revived researchers' interest in applications for medical imaging in recent years. During the past decade, results from numerous published articles have shown the ability of texture analysis algorithms to extract diagnostically meaningful information from medical images that were obtained with various imaging modalities, such as chest radiography, mammography, US, computed tomography (CT), single photon emission CT, positron emission tomography, and magnetic resonance imaging (3–12).

Nonetheless, texture analysis is not a panacea for the diagnostic interpretation of radiologic images. The pursuit of texture analysis is based on the hypothesis that the texture signature of an image is relevant to the diagnostic problem at hand. Furthermore, the effectiveness of texture analysis is bound by the type of algorithm that is used to extract meaningful features. Studies, such as the one by Chen et al (1), are useful because they are performed to test the hypothesis, and their promising results suggest that this type of image analysis might be able to enhance the human perceptional function (particularly in inexperienced observers), when textural features can be comprehensively described.

Putting It All Together

With image texture analysis in the role of the visual perceptional function, the process of feature extraction and image coding is achieved. The extracted features can now be merged into a diagnosis by using a decision-making algorithm, with choices that range from the rule-based models to the traditional statistical analysis to the more popular (and often more successful) artificial intelligence techniques, such as neural networks and genetic algorithms. By using image texture analysis as the preprocessing step in CAD schemes, the input generation process is automated and, therefore, is reproducible and robust.

The importance of automated feature extraction in the diagnostic interpretation of medical images should not be underestimated. Study findings have shown that the accuracy, efficiency, and consistency of any decision algorithm can be seriously compromised by possible bias or variability that is introduced when the input features are generated (28,29). Consequently, an algorithm for image interpretation that is based on visually extracted features could possibly lack consistency. Factors such as attitude, beliefs, preconceptions, expectations, and fatigue can cause differences in image perception among radiologists and, thus, can introduce inter- and intraobserver variability at the input-generation level.

A decision algorithm that uses automatically extracted features has a better chance of producing a robust CAD system. The article by Chen et al (1) serves as a case in point. Findings from a recently published article (30) show that the radiologists' descriptions and diagnostic assessments of sonograms of solid breast masses suffer from considerable observer variability, even though the appearance of masses was described according to a lexicon that was proposed in an earlier benchmark study (31).

A CAD tool that uses automatically extracted features addresses an established clinical weakness of the diagnostic process and also complements the radiologists' perceptive abilities. Texture analysis is well suited to this problem since the radiologists themselves rely on visual texture to detect and describe breast lesions on US images. Furthermore, the autocorrelational function used in this study (1) to capture the textural properties on the images seems to be a justified choice. Other authors have supported the discriminatory ability of the autocorrelational function for US tissue characterization (32). The fact that the image processing used in the study can be easily performed with existing software packages dramatically simplifies the incorporation of the proposed CAD tool into the clinic.

Nonetheless, the selection of textural features will always affect the diagnostic performance of the final CAD scheme. Optimal selection of an adequate set of textural features is a challenge, especially with the limited data we often have to deal with in clinical problems. Consequently, the effectiveness of any CAD tool will always be conditional on two things: (*a*) how well the selected features describe the disease states that need to be discriminated and (*b*) how well the study group reflects the overall target patient population for the CAD tool. Thus far, results of studies such as the one reported by Chen et al (1) are good reasons to believe that narrowly defined clinical applications are amenable to the current status of texture analysis and to computer-aided decision-making methods.

However, it is time to take the next step toward a more sophisticated and productive analysis of CAD research. It is important to evaluate the diagnostic performance of proposed CAD systems in clinical cases with variable difficulty. In addition, robustness should be an important component of the evaluation. Applications on limited numbers of images do not demonstrate robustness and do not permit fair comparisons among similar attempts by other investigators. Thus, the need for unified and publicly available databases of medical images is strong. These databases should be designed to include clinically important cases with various levels of diagnostic difficulty and files of established truths that will allow CAD researchers to evaluate and compare their techniques. We need to remember that in clinical applications emphasis is placed more on the diagnostic performance than on the sophistication of the methods. Methodological complexity is justified only if it is accompanied by substantial diagnostic improvement, as demonstrated in benchmark image databases.

Future Perspectives in CAD Research

CAD analysis has faced skepticism and numerous criticisms in the past. Regardless, CAD research made considerable progress by showing potential in multiple clinical areas and also by producing practical, commercialized applications that have withstood thorough testing and that have gained approval from the U.S. Food and Drug Administration (eg, Image Checker M1000 [R2 Technology, Los Altos, Calif] for screening mammograms). Thus far, study findings have shown that CAD can enhance the diagnostic performance of radiologists, if it is used as a second opinion (33,34). However, it is still difficult to predict its acceptance in medical diagnosis, and its acceptance will not be determined until findings from extensive clinical use demonstrate quantifiable benefits. At the same time, commercial CAD products will have to face new challenges related to marketing, technical support, modifications, and product life-cycles, as the needs and practices in the health care environment are constantly redefined.

Because the trend in radiology is toward increasingly digital technology, computerized management of images is the rule rather than the exception. The evolution in computer technology and in hardware for real-time signal processing is the driving force that will shape the course of CAD as an additional step in image manipulation in radiology. For example, the CAD tool proposed by Chen et al (1) can be easily integrated into the existing hardware that is used for image acquisition to provide a quick second opinion, as figure 5 of their article indicates. If we decide to exploit the well-known advantages of computers (processing speed, almost unlimited memory, consistency, availability), we can design CAD tools to assist with some well-known clinical needs, such as the analysis of difficult-to-interpret images, the short supply of required expertise, and the cost-effective utilization of available resources.

Even with all of the progress in the past decade, CAD is still a long way from fulfilling our vision. The ideal CAD workstation of the future will be an intelligent system that does not need to be painstakingly programmed. It will have the human abilities to transfer acquired knowledge to new tasks, to adapt to the diagnostic problem, to choose image features that are relevant to the clinical task, to analyze the image, to offer diagnostic suggestions, and, finally, to justify the suggestions on the basis of available reference data. That CAD system will be a true partner to the diagnostic radiologist.

It is difficult to speculate about when the scientific and technical developments will turn this exciting prospect into reality. Meanwhile, CAD research will continue to be constrained to narrowly defined diagnostic problems to demonstrate clinical effect. In that context, image texture analysis as a preprocessing step for images plays an important role in the research of CAD, with the ultimate goal of enhancing the radiologists' visual perceptive skills.

References

1. Chen DR, Chang RF, Huang YL. Computer-aided diagnosis applied to US of solid breast nodules by using neural networks. Radiology 1999; 213:407–412.
2. Jackson VP. Management of solid breast nodules: what is the role of sonography? (editorial). Radiology 1995; 196:14–15.
3. Morishita J, Doi K, Katsuragawa S, Monnier-Cholley L, MacMahon H. Computer-aided diagnosis for interstitial infiltrates in chest radiographs: optical-density dependence of texture measures. Med Phys 1995; 22:1515–1522.
4. Vittitoe NF, Baker JA, Floyd CE Jr. Fractal texture analysis in computer-aided diagnosis of solitary pulmonary nodules. Acad Radiol 1997; 4:96–101.
5. Wei D, Chan HP, Helvie MA, et al. Classification of mass and normal breast tissue on digital mammograms: multiresolution texture analysis. Med Phys 1995; 22:1501–1513.
6. Chan HP, Sahiner B, Petrick N, et al. Computerized classification of malignant and benign microcalcifications on mammograms: texture analysis using an artificial neural network. Phys Med Biol 1997; 42:549–567.
7. Petrick N, Chan HP, Wei D, Sahiner B, Helvie MA, Adler DD. Automated detection of breast masses on mammograms using adaptive contrast enhancement and texture classification. Med Phys 1996; 23: 1685–1696.
8. McPherson DD, Aylward PE, Knosp BM, et al. Ultrasound characterization of acute myocardial ischemia by quantitative texture analysis. Ultrason Imaging 1986; 8:227–240.
9. Ito M, Ohki M, Hayashi K, Yamada M, Uetani M, Nakamura T. Trabecular texture analysis of CT images in the relationship with spinal fracture. Radiology 1995; 194: 55–59.
10. Staff RT, Gemmell HG, Duff PM, et al. Decompression illness in sports divers detected with technetium-99m-HMPAO SPECT and texture analysis. J Nucl Med 1996; 37:1154–1158.
11. Lucht R, Brix G, Lorenz WJ. Texture analysis of differently reconstructed PET images. Phys Med Biol 1996; 41:2207–2219.
12. Kjaer L, Ring P, Thomsen C, Henriksen O. Texture analysis in quantitative MR imaging: tissue characterisation of normal brain and intracranial tumours at 1.5 T. Acta Radiol 1995; 36:127–135.
13. Robinson PJ. Radiology's Achilles' heel: error and variation in the interpretation of the Roentgen image. Br J Radiol 1997; 70:1085–1098.
14. Anderson RE, Hill RB, Key CR. The sensitivity and specificity of clinical diagnostics during five decades: toward an understanding of necessary fallibility. JAMA 1989; 261:1610–1617.
15. Renfrew DL, Franken EA Jr, Berbaum KS, Weigelt FH, Abu-Yousef MM. Error in radiology: classification and lessons in 182 cases presented at a problem case conference. Radiology 1992; 183:145–150.
16. Brady AP, Stevenson GW, Stevenson I. Colorectal cancer overlooked at barium enema examination and colonoscopy: continuing perceptual problem. Radiology 1994; 192:373–378.
17. Berlin L. Malpractice issues in radiology: perceptual errors. AJR 1996; 167:587–590.
18. Berlin L, Hendrix RW. Malpractice issues in radiology: perceptual errors and negligence. AJR 1998; 170:863–867.
19. Berlin L. Errors in judgment. AJR 1996; 166:1259–1261.
20. Durgin FH, Proffitt DR. Visual learning in the perception of texture: simple and contingent aftereffects of texture density. Spat Vis 1996; 9:424–474.
21. Folfiak P. Forming sparse representations by local anti-Hebbian learning. Biol Cybern 1990; 64:165–170.
22. Field DJ, Hayes A, Hess RF. Contour integration by the human visual system: evidence for a local "association field." Vision Res 1993; 33:173–193.
23. Karni A, Sagi D. The time course of learning a visual skill. Nature 1993; 365:250–252.

24. Tuceryan M, Jain AK. Texture analysis. In: Chen C, Pau L, Wang P, eds. Handbook of pattern recognition and computer vision. River Edge, NJ: World Scientific, 1993; 235–276.
25. Julesz B, Gilbert EN, Shepp LA, Frish HL. Inability of humans to discriminate between visual features that agree in second-order statistics—revisited. Perception 1973; 2:391–405.
26. Julesz B. Visual discrimination of textures with identical third-order statistics. Biol Cybern 1978; 31:137–140.
27. Harlow CA, Eisenbeis SA. The analysis of radiographic images. IEEE Trans Comput 1973; C22:678–689.
28. Egglin TK, Feinstein AR. Context bias: a problem in diagnostic radiology. JAMA 1996; 276:1752–1755.
29. Scott JA, Fisher RE, Palmer EL. Neural networks in ventilation-perfusion imaging. II. Effects of interpretive variability. Radiology 1996; 198:707–713.
30. Baker JA, Kornguth PJ, Soo MS, Walsh R, Mengoni P. Sonography of solid breast lesions: observer variability of lesion description and assessment. AJR 1999; 172: 1621–1625.
31. Stavros AT, Thickman D, Rapp CL, Dennis MA, Parker SH, Sisney GA. Solid breast nodules: use of sonography to distinguish between benign and malignant lesions. Radiology 1995; 196:123–134.
32. Huisman HJ, Thijssen JM. Adaptive texture feature extraction with application to ultrasonic image analysis. Ultrason Imaging 1998; 20:132–148.
33. Kegelmeyer WP Jr, Pruneda JM, Bourland PD, Hillis AH, Riggs MW, Nipper ML. Computer-aided mammographic screening for spiculated lesions. Radiology 1994; 191:315–317.
34. Jiang YL, Nishikawa RM, Schmidt RA, Metz CE, Giger ML, Doi K. Improving breast cancer diagnosis with computer-aided diagnosis. Acad Radiol 1999; 6: 22–33.