Measuring Pixel Classification Accuracy Using Synthetic Spectrum Images

John Henry J. Scott and Nicholas W. M. Ritchie

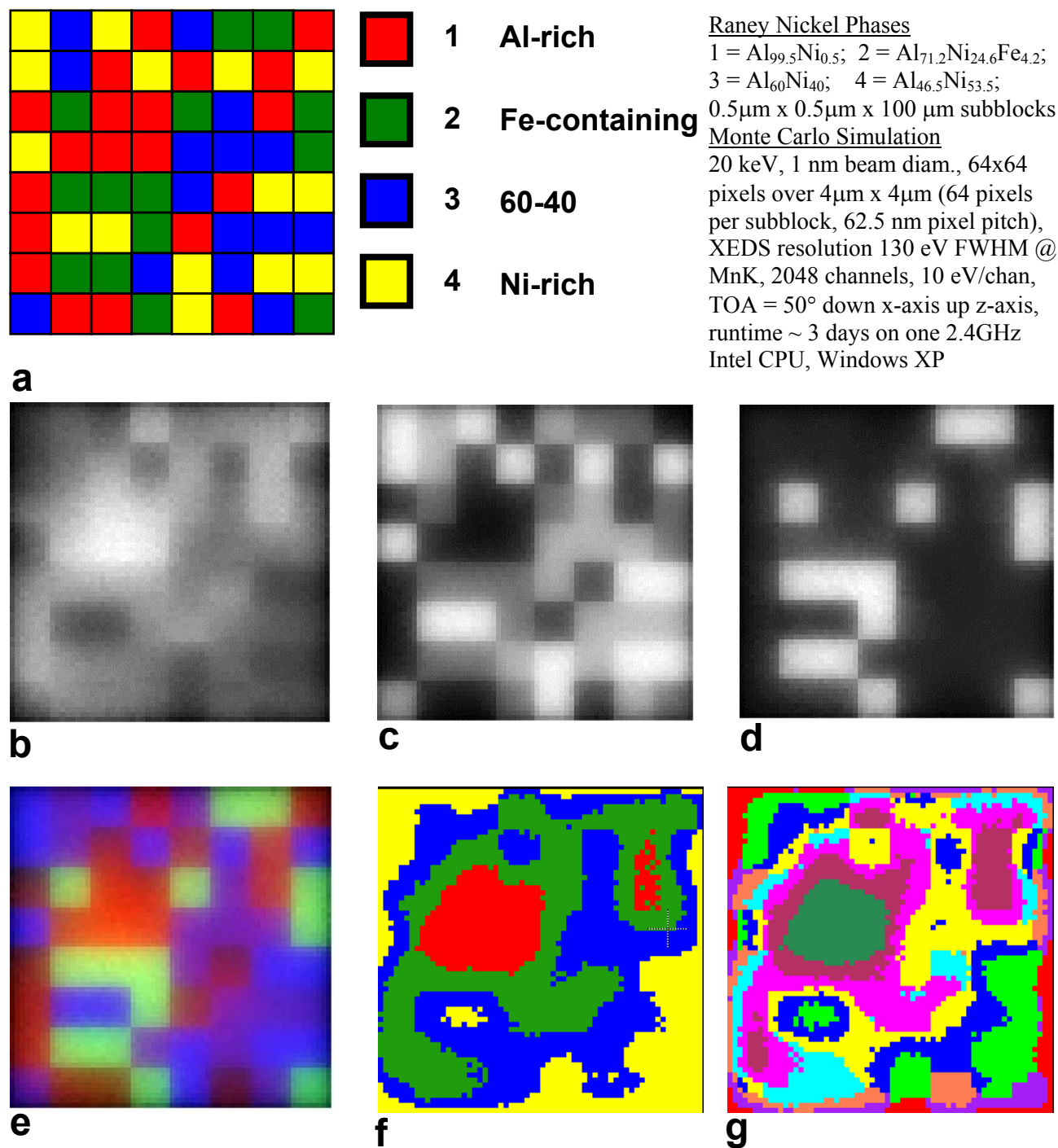National Institute of Standards and Technology, Gaithersburg, MD 20899-8370

As the use of data acquisition hardware capable of acquiring spectrum images becomes widespread, many analysts are employing pixel classification techniques to process their microanalysis datasets. Before the advent of spectrum imaging, analysts classified structural features in their samples using elemental x-ray maps. Today, many commercial microanalysis packages come equipped with some form of classification procedure for assigning pixels or larger spatial features to one of a small number of categories or classes. These class labels are presented to the analyst to provide insight into the chemical heterogeneity of the sample, or sometimes even its spatial distribution of "phases".

This proliferation of classification techniques within the microanalysis community has led to the need for accurate measures of performance of the various classification algorithms. Different pixel classification methods vary in their accuracy, precision, and specificity; they exhibit different strengths and weaknesses depending on the situation and what information is to be extracted from the data. The methods also exhibit different levels of sensitivity to problems in the data such as the presence of noise or badly overlapped spectral peaks. Thus, the goal of post-classification performance metrics is to illuminate how these methods meet the required analytical challenges and how gracefully their performance degrades when applied to poor or unusual data.

The best measures of performance for classifiers of all types are based on comparison of their predicted class assignments with known class membership. With real samples this is nearly impossible since the spectrum image is often the only source of phase information available at the relevant length scales. The calculation of synthetic spectrum images from computer models (Fig. 1a-e) obviates this problem, since both the composition and microstructural geometry are precisely known for the phantom samples [1]. This allows the creation of a "ground truth" image revealing the true class membership of each pixel. Using this ground truth image as a key, the performance of any given classifier can be assessed, including measurement of error rates (both errors of commission and omission). When the classifier returns the same number of classes as the ground truth (Fig. 1f), all errors can be tabulated in a confusion matrix [2]; when the number of classes is variable (Fig. 1g) error assessment is not as straightforward. Classifiers with thresholds as free parameters (such as Bayes classifiers) benefit from analysis of their receiver-operator characteristic (ROC) curves, which reveals the relationship of false negative vs. false positive classification error rates as a function of the cut-off threshold.

References
[1] N.W.M. Ritchie, "A New Monte Carlo Application for Complex Sample Geometries", *Surface and Interface Analysis* **37** (2005) 1006-1011.
[2] T. Hastie, R. Tibshirani, J. Freidman, *The Elements of Statistical Learning*, Springer, New York (2001) pp 461-465.
[3] J. R. Jensen, *Introductory Digital Image Processing*, Prentice-Hall (1996).

	1	Al-rich
	2	Fe-containing
	3	60-40
	4	Ni-rich

Raney Nickel Phases
$1 = Al_{99.5}Ni_{0.5}$; $2 = Al_{71.2}Ni_{24.6}Fe_{4.2}$;
$3 = Al_{60}Ni_{40}$; $4 = Al_{46.5}Ni_{53.5}$;
0.5μm x 0.5μm x 100 μm subblocks
Monte Carlo Simulation
20 keV, 1 nm beam diam., 64x64
pixels over 4μm x 4μm (64 pixels
per subblock, 62.5 nm pixel pitch),
XEDS resolution 130 eV FWHM @
MnK, 2048 channels, 10 eV/chan,
TOA = 50° down x-axis up z-axis,
runtime ~ 3 days on one 2.4GHz
Intel CPU, Windows XP

Fig 1.(a) ground truth image showing arrangement of the four Raney Nickel phases simulated by 3D Monte Carlo; a 64x64 pixel spectrum image (SI) was generated with an XEDS spectrum at each pixel **(b)** Al K x-ray map derived from the SI **(c)** Ni K x-ray map **(d)** Fe K x-ray map **(e)** RGB overlay of the three elemental maps with red=Al, green=Fe, and blue=Ni **(f)** unsupervised pixel classification result using iterative K means [2] and four pixel categories to classify the first 20 principal components of the SI; the poor quality of the classification is deliberate to illustrate the utility of post-classification performance metrics **(g)** unsupervised pixel classification result using the ISODATA iterative algorithm [3] to classify the first 20 principal components; the number of classes is a free parameter in this algorithm, complicating measurement of classification accuracy.