ON-BELT ANALYSIS OF MINERALS USING NATURALLY OCCURRING GAMMA RADIATION*

Bill Moran[1], Du Huynh[2], Michael Edwards[3], Andrew Harris[3], Xuezhi Wang[1], Barbara La Scala[1]

[1]Department of Electrical and Electronic Engineering, The University of Melbourne, Australia
[2]School of Computer Science and Software Engineering, The University of Western Australia, Australia
[3]Scantech, Camden Park SA 5038, Australia

[1]{b.moran, xu.wang, bflscala}@ee.unimelb.edu.au [2]du@csse.uwa.edu.au [3]{m.edwards, a.harris}@scantech.com.au

ABSTRACT

We describe a method to analyze materials on a conveyor belt using natural gamma spectra collected with a BGO (Bismuth Germanate) gamma ray detector, which collects emissions from Potassium (K), Uranium (U), and Thorium (Th) in the materials. A statistical model is proposed based on a Poisson process and an approximate maximum likelihood (ML) technique via the expectation-maximization (EM) algorithm is then used to estimate the amount of each of the three elements in the material. A refinement of the statistical model is used to estimate linear drift in the detector.

Index Terms— Gamma-ray detectors, Poisson processes, Maximum likelihood estimation, Expectation-maximization, Calibration.

1. INTRODUCTION

We describe, and compare with more conventional techniques, a method for analysis of minerals on conveyor belts using naturally occurring gamma ray emissions. The natural gamma analyzer is essentially a BGO detector housed in a temperature controlled lead box beneath a conveyor belt, with a lead shield above the conveyor belt. The lead serves to shield the detector from terrestrial and cosmic gamma radiation which would otherwise add noise to the spectrum. The detector is connected to a digital multi-channel analyzer, which sorts the measured gamma rays into a 1024 channel spectrum by energy. For a coal application, the analyzer provides ash (non-combustible mineral matter) content typically every 2 to 15 minutes without contacting the coal or requiring the addition of sealed radioactive sources. This technique relies on strong linear correlations between concentrations of K, U and Th and ash in coal, and leads to a calibration equation for predicting ash content of coal.

Gamma ray spectra are collected every 900 seconds from a BGO detector with 1024 channels linearly spaced across energies up to 3 MeV. To maintain spectrum stability, the voltage applied to the detector is adjusted at the start of each sample collection. Analysis is based on the prior collection of elemental spectra for K (from potassium chloride), U (from torbenite rock) and Th (from a gas mantle), which we refer to as *library spectra*. These 3 chemical elements are the only ones to emit gamma radiation. Figure 1 shows the library spectra of these 3 elements.

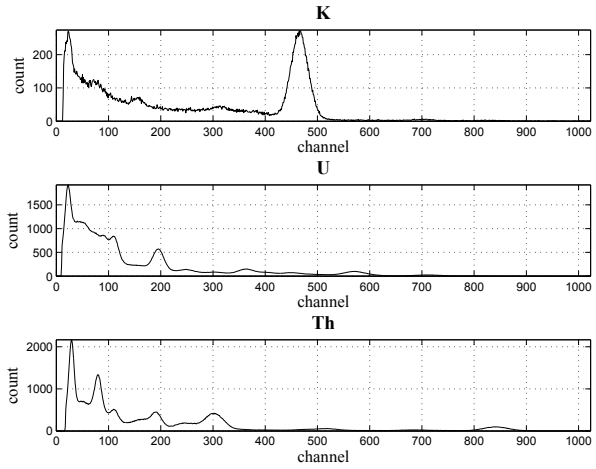


Fig. 1. Library spectra for K, U, and Th. Each spectrum has a very distinctive signature: Spectrum K has a prominent peak in the channel interval [420,520] where spectra U and K are flat; spectra U and Th have unique peaks in the [530,630] and [750,950] intervals.

Our goal is to analyze the amounts of K, U and Th present in the material, and as part of this process, we estimate the detector drift. Conventional calibration methods simply sum the counts in wide windows around each major peak in the spectrum. Slight detector drifts will cause the peaks to move within these windows and so have only a small impact on the analyzer accuracy. Of course, large detector drifts which cause these peaks to move partly or wholly outside the windows cause inaccurate analysis. Because there is no ubiquitous reference peak in the spectrum, automatic adjustment of detector voltage to compensate for detector drift is not nor-

*A provisional patent (Australian Application Number 2007905412) for the algorithm was filed in October 2007.

mally implemented. Over time, the gain of the detector's photo-multiplier tube changes. This is minimized by controlling the detector temperature, but electronics drift is inevitable. The main deficiency of the window approach is that only parts of the spectrum are employed, which increases errors compared with a method which involves consideration of the entire spectrum. A calibration method using the entire spectrum permits the use of a smaller, less expensive detector to achieve the same accuracy.

Here we describe our alternative ML approach using a Poisson model for channel counts. This gives equations which seem to be difficult to solve directly, so we resort to an EM approximation technique. The inclusion of a linear model for drift further complicates the calculations even within the EM approach, and we employ a simpler method for drift estimation using a search for the drift parameters based on the Levenberg-Marquardt variant [1, 4] of gradient descent.

2. PROBLEM FORMULATION

Our initial derivation ignores drift of the spectral peaks over time. The three spectra for K, Th, and U, the only primary elements naturally emitting gamma radiation, as functions of the channels, are denoted by $\mathbf{S}^{\mathrm{K}}, \mathbf{S}^{\mathrm{U}}, \mathbf{S}^{\mathrm{Th}}$ or more conveniently for writing formulae as $S^i(c)$, $i = 1, \cdots, 3,\ c = 1, \cdots, n_c$, where n_c denotes the number of channels, $\mathbf{b}(c)$ the datum; that is, a spectrum from a mineral sample, and $\alpha^{\mathrm{K}}, \alpha^{\mathrm{U}}, \alpha^{\mathrm{Th}}$ the amounts of potassium, uranium and thorium present in the sample. We also use $\boldsymbol{\alpha} = (\alpha^1, \alpha^2, \alpha^3)^\top$ to represent these three amounts when convenient.

A simple (and currently used) approach to the estimation of $\alpha^{\mathrm{K}}, \alpha^{\mathrm{U}}, \alpha^{\mathrm{Th}}$ involves a linear regression:

$$\sum_{i=1}^{3} \alpha^i \mathbf{S}^i = \mathbf{b}. \tag{1}$$

This is, of course, an over-determined system and typically a least-squares approach is taken to solving it. Our approach more closely models the physical processes generating the data. We assume that the number of counts in the c^{th} channel is a Poisson random variable X_c, and that these random variables are independent both between channels and between elements. Let b_c be the number of counts occurring in the c^{th} channel. We calculate the likelihood in a form that will facilitate, at least approximately, ML estimation (see, e.g., [2]) of the parameters.

The likelihood of a given datum $\mathbf{x}$ is then

$$\begin{aligned}\Lambda_{\mathbf{X}}(\boldsymbol{\alpha}) &= P(X_c = b_c \text{ counts in the } c^{\text{th}}\text{channel}, \forall c) \\ &= e^{-\sum_c (\sum_{i=1}^3) \alpha^i S^i(c)} \prod_{c=1}^{n_c} \frac{1}{b_c!} \left(\sum_{i=1}^{3} \alpha^i S^i(c) \right)^{b_c},\end{aligned}$$

which is a product of Poisson distributions with mean $\alpha^{\mathrm{K}} S_c^{\mathrm{K}} + \alpha^{\mathrm{U}} S_c^{\mathrm{U}} + \alpha^{\mathrm{Th}} S_c^{\mathrm{Th}}$. Our justification for independence is as follows: each separate atom gives rise to its own independent decay chain and it is unlikely in the context of large numbers of atoms of these materials being present in the sample that the analyzers will receive gamma rays from different parts of the same decay chain.

Here, $b_c^{\mathrm{K}}, b_c^{\mathrm{U}}, b_c^{\mathrm{Th}}$, for $c = 1, \cdots, n_c$, are variables in the sum of b_c. The direct approach to ML estimation involves differentiation of the log-likelihood with respect to the unknowns and setting to zero. This gives the following equations:

$$\sum_{c=1}^{n_c} S^i(c) = \sum_{c=1}^{n_c} \frac{b_c S^i(c)}{\sum_{j=1}^{3} \alpha^j S^j(c)}. \tag{2}$$

We need to solve these equations for the unknowns α^{K}, α^{U}, α^{Th}. We use an EM approach.

3. EXPECTATION-MAXIMIZATION APPROACH

The EM method approximates the ML solution to an estimation problem. We refer to [3] for a discussion of the EM technique. It works by assuming that there are measurements for certain "hidden variables". The likelihood is computed on the basis of these hidden data and the ML solution obtained. Usually this is more tractable than the original ML problem.

In our case the hidden data are the unknown counts b_c^{K}, b_c^{U}, b_c^{Th}. The distribution of these hidden data is:

$$\begin{aligned}\Lambda_{\mathbf{Y}}(\boldsymbol{\alpha}) &= P(Y_c = (b_c^{\mathrm{K}}, b_c^{\mathrm{U}}, b_c^{\mathrm{Th}}) \text{ counts in } c^{\text{th}}\text{channel}, \forall c) \\ &= \prod_{c=1}^{n_c} \left(e^{-\sum_{i=1}^3 \alpha^i S^i(c)} \frac{\prod_{i=1}^3 (\alpha^i S^i(c))^{b^i(c)}}{\prod_{i=1}^3 b^i(c)!} \right), \end{aligned} \tag{3}$$

so the log-likelihood is

$$\log \Lambda_{\mathbf{Y}}(\boldsymbol{\alpha}) = \sum_{c=1}^{n_c} \sum_{i=1}^{3} \Big(-(\alpha^i S^i(c)) - \log(b^i(c)!) + b^i(c) \log(\alpha^i S^i(c)) \Big).$$

The EM method consists of two steps:

- **The Expectation step** (or the **E-step**). The conditional expectation of the log-likelihood with respect to the actual data $\mathbf{X}$ is computed.
- **The Maximization step** (or the **M-step**). Based on the imputed data obtained from the E-step, the unknown parameters, $\alpha^{\mathrm{K}}, \alpha^{\mathrm{U}}$, and α^{Th} are estimated to maximize the log-likelihood.

The result of one EM step will give an estimate whose likelihood is at least as large as its predecessor [3]. Typically the EM process is iterated several times.

We refrain from giving the details of the calculations involved in this case. They result in an updated value of the α^i of the form:

$$\alpha^i = \sum_{c=1}^{n_c} \left(\frac{{\alpha^i}_0 S^i(c)}{\sum_{i=1}^3 {\alpha^i}_0 S^i(c)} b_c \right) \Bigg/ \sum_{c=1}^{n_c} S_c^{\mathrm{K}}\ . \tag{4}$$

We can merge the E-step and M-step, yielding the following iterative scheme for estimation of $\alpha^{\mathrm{K}}, \alpha^{\mathrm{U}}, \alpha^{\mathrm{Th}}$:

EM-1) Take the counts $(b_c)_{c=1}^{n_c}$;

EM-2) Guess a value for $\alpha_0^{\mathrm{K}}, \alpha_0^{\mathrm{U}}, \alpha_0^{\mathrm{Th}}$ — this might be based on a simple use of estimates of the mean for the Poisson distributions;

EM-3) Use the formulae (4) to iteratively find $\alpha_n^{\mathrm{K}}, \alpha_n^{\mathrm{U}}, \alpha_n^{\mathrm{Th}}$ by replacing $\alpha_0^{\mathrm{K}}, \alpha_0^{\mathrm{U}}, \alpha_0^{\mathrm{Th}}$ by $\alpha_{n-1}^{\mathrm{K}}, \alpha_{n-1}^{\mathrm{U}}, \alpha_{n-1}^{\mathrm{Th}}$;

According to the EM theory this iteration never decreases the likelihood, and typically converges to the ML solution.

4. INCORPORATING DRIFT

At this point we incorporate drift of the detectors into the model. Empirical evidence suggests that, to a reasonable approximation, *drift* is linear in the channels. That is, the data originally going into channel c now enters channel $\beta c + \gamma$, where β and γ are two new unknown parameters. Occurrences of $b^i(c)$ are now substituted by $b^i(\beta c + \gamma)$. Let $\boldsymbol{\alpha}_0 = (\alpha_0^1, \alpha_0^2, \alpha_0^3)^\top$ be the initial estimates of the amounts, and let β_0 and γ_0 be the initial estimates of drift. The log-likelihood function becomes:

$$\log \Lambda_{\mathbf{Y}}(\boldsymbol{\alpha}, \beta, \gamma) = \sum_{i=1}^{3} \sum_{c=1}^{n_c} \Big(-\alpha^i S^i(c) - (\log(b^i(\beta c + \gamma))!) + b^i(\beta c + \gamma) \log(\alpha^i S^i(c)) \Big). \quad (5)$$

EM calculations for the estimation of $\boldsymbol{\alpha}$ yield

$$\alpha^i = \frac{\sum_{c=1}^{n_c} \mathsf{E}[b^i(\beta c + \gamma) \,|\, \mathbf{X}, \boldsymbol{\alpha}_0, \beta_0, \gamma_0]}{\sum_{c=1}^{n_c} S^i(c)}, \quad (6)$$

where

$$\mathsf{E}[b^i(\beta c + \gamma) \,|\, \mathbf{X}, \boldsymbol{\alpha}_0, \beta_0, \gamma_0] = \frac{\alpha_0^i S^i(\tau(c))}{\sum_{j=1}^{3} \alpha_0^j S^j(\tau(c))} b(\tau(c)),$$

and $\tau(c) = \left(\frac{\beta c + \gamma - \gamma_0}{\beta_0} \right)$. Calculation of drift parameters via the EM algorithm is difficult, and we adopt a simplified approach. With the detectors regularly calibrated, the drift involved is quite small, but not insignificant enough to be neglected. We propose a standard optimization approach, the Levenberg-Marquardt (LM) method [1, 4], to iteratively refine the drift parameters β and γ, using the initial estimates $\beta = 1$ and $\gamma = 0$.

Let $\mathbf{d} = (\beta, \gamma)^\top$ and let

$$f(c; \mathbf{d}) = b^i(\beta c + \gamma) - \sum_{i=1}^{3} \alpha^i S^i(c) \quad (7)$$

$$\mathbf{f}(\mathbf{d}) = (f(1; \mathbf{d}), \cdots, f(n_c; \mathbf{d}))^\top. \quad (8)$$

We design the following objective function to be minimized by LM:

$$\min_{\beta, \gamma} F(\beta, \gamma) = \mathbf{f}^\top(\beta, \gamma) \mathbf{f}(\beta, \gamma). \quad (9)$$

The LM optimization process is carried out in an alternating fashion with the EM approach for the estimation of $\boldsymbol{\alpha}$. The entire procedure of our algorithm is thus:

1. Set $\mathbf{d} = (1, 0)^\top$.
2. Use the current estimate $\mathbf{d}$ for the drift parameters, apply the EM algorithm to estimate $\boldsymbol{\alpha}$, i.e., the amounts for K, U, and Th.
3. By using the current estimate of $\boldsymbol{\alpha}$, apply the LM steps above to refine $\mathbf{d}$.
4. Go back to step 2.

We have found that the above procedure need only be iterated 4 times to obtain good alignments of the peaks of the spectra. We call this algorithm of estimating $\boldsymbol{\alpha}$ and $\mathbf{d}$ in an alternate fashion the EM-LM algorithm.

5. RESULTS

An independent analysis is not readily available to produce "true" values for the $\boldsymbol{\alpha}$ parameters or for the drift parameters for our real data. As a result, we have tested the algorithm in the following two ways:

1. Application to synthetically generated data based on the library spectra and prescribed values for the $\boldsymbol{\alpha}$, β, and γ parameters.
2. Application to collections of real data, where performance is measured by fidelity of the reconstructed spectra using the estimated parameters.

For each test on synthetic data, a random $\boldsymbol{\alpha}$ vector was first generated with values of α^i around the range expected in the minerals; the three hidden variables b_c^{K}, b_c^{U}, and b_c^{Th}, for each channel c, were then synthesized as a Poisson Process with mean equal to S_c^{K}, S_c^{U}, and S_c^{Th}. The total values of these three hidden variables were then summed to give the count $b(c)$, giving a test spectrum that is free of drift. The drift parameters β and γ were finally synthesized randomly and the drift affected spectrum was generated via spline interpolation. For the testing on real data, the channel counts of the spectra vary greatly depending on the duration the materials were under the BGO detectors and the composition of the materials.

We compared the accuracy of the EM approach for estimating $\boldsymbol{\alpha}$, with the technique which involves replacement of the EM step (i.e., step 2) of the EM-LM algorithm by the linear least squares solution given by (1). We term this algorithm the LLS-LM algorithm.

We present results from 3 sets of 100 tests with synthetic data, simulating the cases of high, medium and low channel counts. In each test, the spectrum was passed to the EM-LM and LLS-LM algorithms. To compare the performances of the two algorithms, the following error measures were computed: the mean percentage errors of $\boldsymbol{\alpha}$ and $\mathbf{d}$ from the known ground truth values in the simulation.

		$\epsilon_{\boldsymbol{\alpha}}$	$\epsilon_{\mathbf{d}}$
Set 1	LLS-LM	(0.79%,1.33%,1.64%)	(0.03%,2.06%)
	EM-LM	(0.40%,0.40%,0.42%)	(0.01%,1.20%)
Set 2	LLS-LM	(0.90%,1.44%,1.65%)	(0.04%,2.56%)
	EM-LM	(0.45%,0.46%,0.42%)	(0.02%,1.51%)
Set 3	LLS-LM	(1.26%,1.52%,1.76%)	(0.04%,2.59%)
	EM-LM	(0.97%,0.69%,0.73%)	(0.03%,2.65%)

Table 1. Comparison of the LLS-LM and EM-LM algorithms. $\epsilon_{\boldsymbol{\alpha}}$ and $\epsilon_{\mathbf{d}}$ denote the percentage errors of the $\boldsymbol{\alpha}$ and drift values for K, U, and Th.

Table 1 shows the mean percentage errors of $\boldsymbol{\alpha}$ and $\mathbf{d}$ from the two algorithms. It is evident from Table 1 that EM-LM outperforms the LLS-LM algorithm for all cases. In addition, use of LLS to estimate the amounts $\boldsymbol{\alpha}$ when these entities are small often leads to *negative mounts*, which is non-physical.

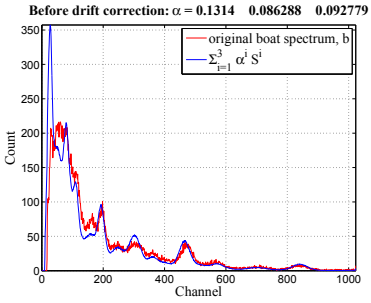
The EM-LM algorithm was applied to many real spectra varying from coal to iron ore. Only one of the results is presented here; in this case the input spectrum is from an iron ore dataset. In Figure 2(a), the blue curve shows the input spectrum from this dataset and the red curve shows the reconstructed spectrum using the library spectra $\{\mathbf{S}^i \mid i = 1, \cdots, 3\}$ and the amounts $\boldsymbol{\alpha}$ estimated from EM without taking drift into account. The two curves are clearly misaligned. Figure 2(b) shows a much better reconstructed spectrum after drift of the detector is corrected using EM-LM. Figure 2(c) and (d) show the input and drift corrected spectra and their differences over the channels.

As mentioned before, it is impossible to obtain the ground truth of $\boldsymbol{\alpha}$. Neither is it possible to obtain the ground truth of the drift parameters, $\mathbf{d}$. However, our experiments confirm that the reconstructed spectra coincide much better with the library spectra when drift correction is incorporated into the algorithm. From our tests on synthetic data, the EM-LM outperforms the linear least squares method.

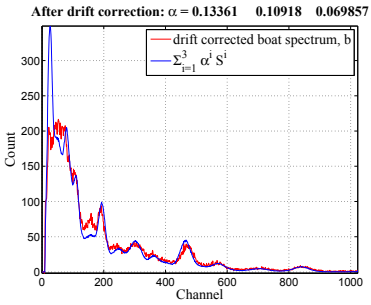
6. CONCLUSION

We have described an estimation method for on-belt analysis of materials using their natural gamma spectra. The technique involves the use of a Poisson model for the generation of the spectra, based on the experimentally determined library spectra for each of these elements. We have incorporated parameters in the model to describe the drift of the detector performance over time. We have described the ML Estimation problem for all of the parameters, its approximation using an EM approach, and finally a technique in which EM is used to estimate the amount of each element present, but employs a more direct approach to the estimation of the detector drift parameters.

The last approach has been implemented on both synthetically generated and real data and been shown, insofar as tests



(a)



(b)

Fig. 2. Output of the EM-LM algorithm on the EM020911 spectrum from an iron ore dataset. (a) Input and reconstructed spectra, $\mathbf{b}$ and $\tilde{\mathbf{b}}$, without drift correction (using EM only); (b) Input and reconstructed spectra, $\mathbf{b}$ and $\hat{\mathbf{b}}$, with drift correction (using EM-LM).

available to us will permit, to be significantly superior to the more conventional approach using a linear regression to estimate the amount of each material present.

While we believe that the results obtained here represent a significant advance on previous techniques used in this form of analysis, this paper is a report on ongoing work. We are continuing to investigate alternative and better methods for the overall estimation process associated with on-belt gamma ray spectrometric methods of elemental analysis and this work will be reported elsewhere.

7. REFERENCES

[1] P. E. Gill, W. Murray, and M. H. Wright. *Practical Optimization*. Academic Press, 1981.

[2] S. M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*, volume I. Prentice-Hall PTR, 1993.

[3] T. K. Moon. The Expectation-Maximization Algorithm. *IEEE Signal Processing Magazine*, 13(6):47–60, 1996.

[4] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C: the Art of Scientific Computing*. Cambridge University Press, 1988.