Data management and visualization of x-ray diffraction spectra from thin film ternary composition spreads

I. Takeuchi, C. J. Long, O. O. Famodu, M. Murakami, J. Hattrick-Simpers, and G. W. Rubloff

Department of Materials Science and Engineering and Center for Superconducting Research, University of Maryland, College Park, Maryland 20742

M. Stukowski and K. Rajan
Department of Materials Science and Engineering, Rensselaer Polytechnic Institute, Troy, New York 12180

(Received 9 December 2004; accepted 13 April 2005; published online 1 June 2005)

We discuss techniques for managing and visualizing x-ray diffraction spectrum data for thin film composition spreads which map large fractions of ternary compositional phase diagrams. An in-house x-ray microdiffractometer is used to obtain spectra from over 500 different compositions on an individual spread. The MATLAB software is used to quickly organize the data and create various plots from which one can quickly grasp different information regarding structural and phase changes across the composition spreads. Such exercises are valuable in rapidly assessing the “overall” picture of the structural evolution across phase diagrams before focusing in on specific composition regions for detailed structural analysis. We have also shown that simple linear correlation analysis of the x-ray diffraction peak information (position, intensity and full width at half maximum) and physical properties such as magnetization can be used to obtain insight about the physical properties. © 2005 American Institute of Physics. [DOI: 10.1063/1.1927079]

I. INTRODUCTION

The combinatorial approach to materials has proven to be effective in uncovering new materials phases with enhanced physical properties as well as rapidly mapping composition-structure-property relationships in complex materials systems. In particular, composition spreads experiments are instrumental in mapping compositional phase diagrams of multicomponent systems. Mapping phase diagrams is central in obtaining comprehensive pictures of materials systems, and mapping active physical properties such as magnetism and ferroelectricity as a function of compositional trends which closely resemble or mirror those of bulk counterparts. There are also instances where it has been shown in many systems that one can indeed observe a variety of physical properties that can be mapped for combinatorial libraries and composition spreads with relatively quick turn around. One piece of information of paramount importance is the phase and crystal structure distribution. There are scanning x-ray microdiffraction techniques, which can be employed to obtain a large number of diffraction diagrams from combinatorial samples. Synchrotron x-ray microdiffraction is a natural technique for this task because the high intensity x-ray beam available can be used to quickly step through a large number of positions on a combinatorial library with minimal time to obtain a diffraction spectrum at each point. In-house microdiffraction x-ray units require much longer time to take data per point, but their advantage is that one can perform the experiments locally.

Obtaining and analyzing x-ray diffraction (XRD) spectra for phase and structure determination is a central part of any materials research exercise, and it should play a significant role in combinatorial materials research as well. Ideally, one would take and study x-ray spectra of all the positions on every sample in order to obtain complete mapping of phase and structural information across the combinatorial samples. However, despite the availability of the microdiffraction techniques, there are practical problems one immediately encounters in implementing such an exercise. First of all, one needs to come up with an efficient way to file and store a large number of x-ray spectra data. The sheer quantity of data produced by these processes presents a significant logis-
tical challenge. A common analysis method for a diffraction spectrum is to fit the peaks with reflections of known compounds using software such as JADE by Materials Data, Inc. and TOPAZ by Bruker. Although some software can be automated to fit a large number of x-ray spectra serially, it is a very cumbersome process and fitting does not necessarily produce the “correct answers.” Because the very premise of the combinatorial experimentation is that there are potentially entirely new phases with previously unknown crystal structures in the library, any simple fitting procedure using reflections from known phases is not sufficient. The situation is further complicated by the fact that in any compositional phase diagram, there can be multi-phase regions.

Despite such difficulties, XRD analysis remains a most fundamental method in materials science, and one must find a way to effectively incorporate it into the overall combinatorial strategy. In this article, we discuss our methodology and procedures for data management and visualization of x-ray microdiffraction spectra of combinatorial samples. The ultimate goal of such an exercise is to obtain a comprehensive and accurate mapping of phase and structure distribution across composition spreads of rich and complex materials systems containing previously unknown materials phases. Although the techniques described here only represent a first step toward this goal, we demonstrate that even simple data visualization and analysis exercises can be effective in extracting important insight about the properties of materials systems under study.

II. EXPERIMENT

As a model system, we have looked at the Ni–Al–Mn ternary system. This system contains Heusler alloys that are ferromagnetic shape memory alloys.\textsuperscript{19} We have previously used a scanning superconducting quantum interference device (SQUID) microscope and micromachined arrays of cantilevers to map composition regions which are ferromagnetic and reversible martensites in this system, respectively.\textsuperscript{20} Despite a number of previous phase mapping studies including ours, much is still unknown about different regions of this phase diagram. The Heusler alloys are known to display significant variation in magnetism and martensite properties depending on the chemical order parameter which in turn strongly depends on the processing conditions. It is also well known that many compositions in this system display modulated structures giving rise to extra diffraction peaks.\textsuperscript{21} The present work therefore is a part of an ongoing effort to better understand this ternary system.

Natural thin film composition spreads of the Ni–Mn–Al system were deposited using an ultrahigh-vacuum three gun magnetron co-sputtering system with a base pressure of 10\textsuperscript{\textcontra} Torr (10\textsuperscript{\textcontra} Pa) on 3 in. (76.2 mm) diameter (100) oriented Si wafers. The details of the synthesis procedure can be found in Refs. 7 and 20.

The composition spread wafers were deposited at room temperature followed by an in situ annealing for 2 h in high vacuum at temperatures in the range of 853–970 K. The total processing time (i.e., deposition and heat treatment of the composition spread) of a sample is roughly 5–6 h before it is ready for rapid characterization. After the deposition, the composition of the wafer spread is immediately determined via wavelength dispersive spectroscopy (WDS) in atomic percent. This measurement can determine the percent fraction of each atom contained at each point on the wafer accurately. The total scan time for this measurement is \textlessthan \textapprox 12–14 h for the entire wafer. Figure 1(a) shows a schematic procedure for the synthesis of a ternary composition spread which covers most of the phase diagram.

X-ray microdiffraction of the fabricated films was performed using the \omega-scan mode of a D8 DISCOVER for combinatorial screening (Bruker-AXS). It is equipped with a GADDS two-dimensional (2D) detector which captures the data for a fixed range of 2\theta and \omega at once. Each composition spread wafer contained 535 individual 1.75 mm \times 1.75 mm squares with continuously changing compositions [Fig. 1(b)]. In order to scan the 2\theta range of interest (2\theta from 25 to 75\degree), microdiffraction was performed in three frames for each square. We use an x-ray beam spot size of 500 \mu m in diameter. We have determined that to collect sufficient data, it takes at least 5 min for each square per frame. Even though scanning of the beam is automated, one still needs to preset the positions of all the squares manually. This process can take up to roughly 2 h for an entire wafer with 535 squares. Because we can only scan the entire library for the same frame before moving to the next frame, the entire spread library must be scanned three times to cover the 2\theta range from 20 to 75\degree. For each x-ray microdiffraction frame,
the scan takes \( \approx 45 \text{ h} \) (\( \approx 1.86 \text{ days} \)). Therefore, to complete all three frames takes \( \approx 134 \text{ h} \) (\( \approx 5.57 \text{ days} \)). Once this is accomplished, the microdiffraction data are in the form of 2D detector images. The size of data generated by this process is on the order of 4 gigabytes for a single spread wafer.

Once the actual scans are completed, the raw data are compiled and integrated to obtain the 2\( \theta \) angles and peak intensities using the D8 GADDS program and a script to automate the process. At this stage, the data are converted to numeric two column text data in an ASCII format. In this format, data

FIG. 2. (Color) Example of an x-ray microdiffraction of a binary composition spread. The composition continuously changes from ZnO to MgO. Compositions up to Mg_{0.81}Zn_{0.19}O are plotted. The phase distribution and evolution are readily identified from such a plot. The inset is a schematic of the binary spread (Ref 5).

FIG. 3. (Color) The Spectrum Scroller visualization scheme. X-ray diffraction spectra from every single spot on the spread wafer are organized in such a way so that one can scroll through the compositional phase diagram viewing the spectra one by one or in a scrolling manner. [Animation is available as an EPAPS movie file (see Ref. 35). It is also available at www.combi.umd.edu].
for each square on the composition spread takes up about 45 kbytes, and the total data volume is ~30 MB. This step takes ~7 h. Then the data for each of the three frames for each composition square are merged using a MATLAB script.

III. DATA VISUALIZATION

We have chosen MATLAB as the primary program for the organization, storage and analysis of this large quantity of data because it has the capacity to import data from a variety of sources, store the data in arbitrarily large data structures, manipulate that data using a large base of both built-in and user defined functions, and because it has a powerful set of graphical rendering tools.

It takes about 1 h to import and consolidate all of the relevant data into MATLAB using scripts developed specifically for this purpose. One script consolidates the XRD data from about 1500 ASCII files (three frames for each of 535 compositions) into numerical matrices within a MATLAB workspace. Other scripts are used to import composition data obtained using WDS and quantitative values of other physical properties measured using various rapid characterization techniques. These properties may include resistivity, saturation magnetization, and coercive field. Once the data are imported into MATLAB, one can perform numerical operations and create graphical representations.

Visual inspection of x-ray spectra is a crucial part of the materials analysis process. It is often the case that one gets so accustomed to looking at x-ray spectra of a particular materials system under study that after a while, just a quick glance at a spectrum can suffice to provide most of the essential information about the structural properties of the material. Thus, it is imperative that we are able to view together all of the spectra we take. In composition spread experiments where the composition continuously changes from a known composition region with a known structure to a new region of interest, visual inspection and simple comparison of the spectra can be used to quickly determine evolution in the composition-structure relationship. This has been well illustrated in the binary composition spreads where the composition changes continuously from one to another. In binary systems, three axes are sufficient to render all of the data; one axis describes the composition, one describes the diffraction angle and a third describes the diffraction intensity. The data are then rendered as a surface in that parameter space.

Figure 2 shows XRD data from a Mg\textsubscript{3}Zn\textsubscript{1−x}O composition spread where the composition changes continuously from pure ZnO to MgO. Its synthesis procedure is described in Ref. 5. From such a plot, one can readily discern all the major features. Careful scrutiny of this plot was instrumental in identifying a new epitaxial relationship between coexisting cubic and hexagonal phases of Mg\textsubscript{3}Zn\textsubscript{1−x}O. In ternary systems, an additional axis is needed to describe all of the possible compositions. One immediate solution is to reduce the amount of information. For instance, one can create a plot describing the intensity, angular position, or the full width at half maximum of a limited number of peaks, each representing a different phase. This may suffice for materials systems where there is minimal change in the structure across the spread and where one can track the structural change in terms of the shift in one or two main diffraction peaks. However, this is not the case for complicated systems such as Ni–Mn–Al. In addition to the main reflections, there may be satellite peaks arising from the chemical ordering as well as modulation in the structure. Thus, it is desirable to be able to visualize the entire spectrum for the entire composition spread.

One way to represent this data is as a scalar volumetric array. In this case, two axes (X and Y) are used for the composition, one (Z) is used for the diffraction angle, and the...
Intensity is represented as a scalar at all points in the parameter space. Volumetric scalar arrays of data are common in many other fields, including computed tomography (CT), magnetic resonance imaging, and computer modeling using finite element analysis, and much work has been done on the visualization of these arrays. However, due to the nature of the information contained within an XRD spectrum, not all of the techniques which have been developed are useful for our purposes. To date, there are basically three different techniques used to visualize volumetric data: (1) looking at planar cross sections of the data, (2) the rendering of binary surfaces at some threshold value (isosurfacing) and (3) treating the data as a solid object with varying transparency and/or color values at each voxel, then casting parallel “rays” through it onto some observation plane. The first technique is useful and will be discussed below. The second technique is less useful because it does not present all (or even most) of the important information that is contained in an XRD spectrum. The information that is valuable about an XRD peak is its intensity, diffraction angle and full width at half maximum. Rendering an isosurface at some intensity value throws away all of the information about how the intensity changes as a function of diffraction angle and composition except that it is higher on one side of the surface than the other. This means that it disregards information about the maximum intensity of XRD peaks and the full width at half maximum of these peaks. The third technique may be useful, and may be pursued in the future. In the following paragraphs, we will discuss our adaptations of these techniques to the rendering of volumetric XRD data.

The first scheme, the Spectrum Scroller, allows a user to rapidly view hundreds of XRD spectra and the corresponding compositions in a matter of minutes. The ordering of the spectrums is such that the composition changes only slightly from spectrum to spectrum. This allows the user to see trends in the change of the spectrum as a function of composition. The user interface for this technique is shown in Fig. 3. The graph at the top shows the XRD spectrum. The graph at bottom left shows where the composition for this spectrum is located on the ternary composition diagram. All of the points for which spectra have been obtained are shown in red and the composition for the displayed spectrum is circled in blue. The program can be operated to “scroll” through all of the compositions in a predetermined order. As the program scrolls, the spectrum continuously changes to reflect the structure of the composition indicated by the graph on the bottom left. (A movie animation illustrating the utility of this scheme is available as supplementary information. It is also available at www.combi.umd.edu.) Because the Spectrum Scroller allows one to look at individual spectra in the standard format, it is helpful when one needs to quickly grasp an idea about the distribution of peak patterns across the phase diagram. The act of scrolling enables one to track the evolution in the spectrum as a function of composition. For instance, from this exercise, we were able to quickly identify the regions in the phase diagram where the (220) peak of the austenite phase near 44° splits into the peaks of the lower symmetry martensite phase due to the martensitic transition of the shape memory alloy.

The second visualization technique, the Angle Scroller, allows a user to look at the intensity data for all of the spectrums simultaneously, but it looks at only one angle at a time. The angles are scrolled through in time to get a full picture of the data. The user interface for this technique is shown in Fig. 4. This technique is useful for two main features: (1) it allows one to see in a single image the composition range

![Diagram](image-url)
over which a particular peak “exists” at a particular angle and (2) it allows the user to track the shift in the position of the peak (as a function of angle) in the composition space. The utility of this visualization scheme is illustrated in an animation where the composition region with high diffraction intensity is seen to change continuously as a function of angle (through time). (The animation is available as supplementary information. It is also available at www.combi.umd.edu). We found this scheme to be useful in rapidly assessing the size of the composition region with the same structure as well as for closely monitoring how a particular peak shifts its position in angle as the composition changes. It is worth noting that this method is identical to that of viewing planar cross sections of CT data. In CT, it is often useful to look at cross sections in other planar directions. However, when looking at XRD data of ternary compositions, we are mostly interested in how the diffraction intensity changes with respect to composition. The only planes in which the variation of the data is due to changing composition are those that are either at a constant diffraction angle, or those which are parallel to the diffraction angle axis. In this model, we have chosen to use the planes describing a single angle (and all compositions) because it allows us to look at the variation of data over the entire composition region. If a plane parallel to the diffraction angle axis were to be chosen, then that data would represent a binary (two composition) spread and could best be visualized using the binary visualization technique already discussed.

Both of the preceding schemes have required the use of time as an extra dimension in the visualization of ternary XRD data. The next scheme makes use of three-dimensional (3D) rendering in order to show all the XRD peaks for all compositions simultaneously. This is the Peak Plotter, and it is shown in Fig. 5. In order to create this plot, it is necessary to separately record the positions and the intensity of all the peaks for each composition.

Two techniques have been developed to perform this task. In the first technique, we use a peak search program (written in Igor), which takes a spectrum and creates a list of peaks with their positions and the intensity. This process can be automated and applied to an arbitrary large number of spectra to create a table containing all the peaks. The problem with this technique is that the automated peak search often fails to find small peaks, properly separate adjacent peaks, and/or correctly index the same peak appearing in different spectra as the same peak. It is particularly difficult to handle a situation where a peak evolves into twin peaks as the composition is changed. Since part of our inquiry is to find reversible structural transitions, the problem of not being able to discern divergent peaks is particularly problematic.

As an alternative to the automated peak search approach, a more reliable, semiautomated/semimanual method has been developed. In this approach, a person sifts through all the spectra to record the positions and the intensity of each XRD peak maximum. If done entirely by hand, this would require a significant amount of work. For example, in the Ni–Mn–Al system, there may be as many as 18 peaks in the 2θ range of 25–75° for each composition sampled. Generally, from the 535 spectra, the positions and the intensity of about 3500 peaks are recorded. To reduce the difficulty of the recording task, a program similar to the Spectrum Scroller described above is used. By using this program, the recording and organization of the position and the intensity of the peaks are reduced to a point and click process, instead of being carried out entirely by hand. The task of going through 535 spectra and picking out the 3500 peaks takes about 3 h. Once that is complete, it is possible to make a plot such as shown in Fig. 5. However, if a plot of, say, only the angular position of the peak corresponding to the main phase is desired, then another sorting step is required. The peaks for each spectrum must be indexed such that “the same” peak can be picked out in all of the spectra where it exists. This is accomplished by fitting a surface (usually a plane) between each group of points in Fig. 5. Peaks are then sorted and indexed by assigning a peak number according to which planes they fall between.

In the Peak Plotter scheme, the presence of a peak at a particular 2θ angle and at a particular composition is signified by a point in the 3D space [Fig. 5(a)]. The color of the point can be chosen to represent either the intensity or the full width at half maximum. However, given the fact that there are potentially a large number of peaks which are yet to
be identified reflections from yet to be identified phases, comparing absolute intensity or full width at half maximum is not necessarily useful. In fact, we have found that just marking the peak position is effective in conveying the breadth of structural variation across the phase diagram. Once one decides to zoom into a particular 2θ range, the intensity information can be added [Fig. 5(b)]. In such a plot, shifts in the position and the intensity of selected peaks can be mapped across the entire spread. Since this is plotted in a 3D environment, it is best to see it from multiple perspectives. An animation of a rotating perspective is available as supplementary information. It is also available at www.combi.umd.edu.

The shifts in the peak position and the intensity are often directly correlated with change in the physical properties. Thus, one can use the information in the table of all the peak positions and intensity compiled for this scheme to look for correlation between the structural change and physical properties. One way to find these correlations is to calculate the correlation coefficients between all the columns in the table together with added columns for the values of some physical properties. Correlation coefficients which are near 1 indicate high linear correlation and those near −1 indicate that two parameters are linearly anti-correlated. Linear anti-correlation means that when one parameter has high values relative to its average, the other parameter has low values relative to its average.

Figure 6(a) shows the in-plane remanent magnetization mapping of the Ni–Mn–Al system. There is a large ferromagnetic region near the Ni deficient Mn–Al region of the phase diagram which had not been previously reported. Mn–Al and some Mn–Al–Ni are known as ferromagnets, but to the best of our knowledge, there have not been reports of ferromagnetism in the composition region with such high Al concentration. We have calculated the correlation coefficients between the remanent magnetization and all of the XRD peak information and found that there is a strong anti-correlation between the magnetization and the angular position of the peak which is centered at around 44.5°. This peak is the (220) or similar reflections of the main phase present throughout the phase diagram. The shift in this peak position is mapped in Fig. 6(b). The comparison of the figures indicates that magnetization increases when the lattice parameter increases (the peak shifts toward a lower angle). Mn has a magnetic exchange integral which changes its sign from negative (antiferromagnetic) to positive (ferromagnetic) ordering as the distance between the atoms is increased. From the correlation, therefore, we conclude that even in such an Al-rich region, it is the Mn atoms in the crystal that are giving rise to the ferromagnetism. Thus, we have demonstrated the utility of the XRD analysis for the entire composition spread in obtaining an important insight about the physical property of the material system. Note that this was carried out without in-depth structural analysis involving fitting of the spectra and figuring out the exact crystal structure. In the future, full width at half maximum information for all the XRD peaks can also be included for such analyses.

It is possible to investigate the structure property relation in a broader context. Correlation between structure information and a variety of physical properties can be investigated using the visualization techniques discussed here. Some of the properties of interest which we can map in a high throughput manner include band gap, conductivity, dielectric constant, martensitic transformation temperature, saturation magnetization and coercive field. All of this information is then mapped out to produce a complete phase diagram.

ACKNOWLEDGMENTS

This work was supported by ONR N000140110761 and N00140410085, NSF DMR 0231291, and MRSEC DMR-00-80008.

3 I. Takeuchi, R. B. van Dover, and H. Koinuma, MRS Bull. 27, 301 (2002).
35 See EPAPS Document No. ERSINAK-76-257506 for animation of the figures as movie files. This document can be reached through a direct link in the online article’s HTML reference section or via the EPAPS homepage (http://www.aip.org/pubservs/epaps.html).