



Automated Extraction of Data from Binary Phase Diagrams for Discovery of Metallic Glasses

Bhargava Urala Kota¹(✉), Rathin Radhakrishnan Nair¹, Srirangaraj Setlur¹,
Aparajita Dasgupta², Scott Broderick², Venu Govindaraju¹,
and Krishna Rajan²

- ¹ Department of Computer Science and Engineering, University at Buffalo,
State University of New York, Buffalo, NY, USA
{buralako,rathinra,setlur,govind}@buffalo.edu
- ² Department of Materials Design and Innovation, University at Buffalo,
State University of New York, Buffalo, NY, USA
{adasgupt,scottbro,krajan3}@buffalo.edu

Abstract. We present a study on automated analysis of phase diagrams that attempts to lay the groundwork for a large-scale, indexable, digitized database of phases at different thermodynamic conditions and compositions for a wide variety of materials. For this work, we concentrate on approximately 80 thermodynamic phase diagrams of binary metallic alloy systems which give phase information of multi-component systems at varied temperatures and mixture ratios. We use image processing techniques to isolate phase boundaries and subsequently extract areas of the same phase. Simultaneously, document analysis techniques are employed to recognize and group the text used to label the phases; text present along the axes is identified so as to map image coordinates (x, y) to physical coordinates. Labels of unlabeled phases are inferred using standard rules. Once a phase diagram is thus digitized we are able to provide the phase of all materials present in our database at any given temperature and alloy mixture ratio. Using the digitized data, more complex queries may also be supported in the future. We evaluate our system by measuring the correctness of labeling of phase regions and obtain an accuracy of about 94%. Our work was then used to detect eutectic points and angles on the contour graphs which are important for some material design strategies, which aided in identifying 38 previously unexplored metallic glass forming compounds - an active topic of research in materials sciences.

1 Introduction

Traditionally, document based information retrieval systems have focused on using data from text and, to a lesser extent, from images. They do not extract, analyze or index the content in document graphics (non-pictorial images in articles). Scientific documents often present important information via graphics and

little work has been done in the document analysis community to address this gap. The graphics present in documents are predominantly in the form of line, scatter plots, bar charts, etc. [1]. Most current techniques for interacting with graphics in documents involve user provided metadata. Information graphics are a valuable knowledge resource that should be retrievable from a digital library and graphics should be taken into account when summarizing a multimodal document for indexing and retrieval [2]. Automated analysis of graphics in documents can facilitate comprehensive document image analysis, and the information gathered can support the evidence obtained from the text data and allow for inferences and analysis that would not have otherwise been possible [3]. In this work, the primary focus is on analyzing and interpreting information contained in phase diagrams which are critical for design within the Materials Science and Engineering community.

Phase diagrams serve as a mapping of phase stability in the context of extrinsic variables such as chemical composition with respect to temperature and/or pressure and therefore provide the equilibrium phase compositions and ratios under variable thermodynamic conditions. The geometrical characteristics of phase diagrams, including the shape of phase boundaries and positions of phase boundary junctions have fundamental thermodynamic origins. Hence they serve as a visual signature of the nature of thermo-chemical properties of alloys. The design of alloys for instance, relies on inspection of many such documented phase diagrams and this is usually a manual process. Our objective is to develop an automated document recognition tool that can process large quantities of phase diagrams in order to support user queries which, in turn, facilitate the simultaneous screening of a large number of materials without loss of information.

Further, from the phase diagram images, we readily identify specific types of phase boundary junctions, known as ‘eutectic points’. We have used this test case to show that we can characterize the shape of eutectic points, and provide a meaning to the term ‘deep eutectic’. Deep eutectics are known to be critical for the formation of metallic glasses (i.e. metallic systems without crystalline order), although previously no clear meaning of deep eutectic had been defined in terms of identifying new compounds [4].

Phase diagrams need specific attention primarily because of the way the information is embedded into the diagram. The lines in a phase diagram are not of a continuously changing value like in a line plot, but instead represent a boundary. A phase diagram cannot be expressed by a simple table like most line plots, bar charts etc. Further, text can appear in different orientations and subsequently associating the text with the phase regions (and sometimes vertical lines) is an added complexity that is non-trivial and essential to the final interpretation by materials science domain experts. These characteristics underline the necessity for having a targeted approach to handling this particular class of diagrams.

For our study, we randomly select a small subset of phase diagrams of binary metallic alloy systems where the X-axis is molar fraction percentage and the Y axis is temperature. The goal of our study is to create a database, where given a temperature value and molar fraction percentage for a particular alloy, the

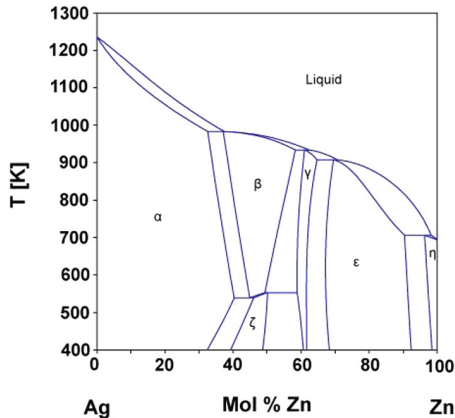


Fig. 1. A typical phase diagram with labeled boundaries.

database returns the phases of the alloy. A typical phase diagram is shown in Fig. 1. Since there are potentially infinite real-valued query points, we evaluate our system on how correctly an entire phase region, i.e., the set of all possible points with the same phase, is labeled.

The rest of the paper is organized as follows: Sect. 2 provides an overview of the related work done for understanding information graphics. Section 3 describes phase diagrams and Sects. 4 through 6 discuss the proposed approach, followed by details of the evaluation metrics and a discussion in Sect. 7.

2 Background

Although graphics analysis and understanding of a variety of diagrams has been addressed in the literature, to the best of our knowledge, no prior work has tackled the problem of analyzing and understanding classes of diagrams with complex semantic interpretation such as phase diagrams. The purpose of graphics in most cases is to display data, including the ones in popular media and research documents. Specifically, in research documents, they serve the purpose of pictorially comparing performance of multiple approaches, and offering objective evaluations of the method proposed in the manuscript. In this section, we discuss some prior work in graphics analysis and understanding. We are primarily interested in analyzing the structure of the graphic and analyzing it to interpret the information present in it. A survey of some of the earliest work in the field of diagram recognition is mentioned in [5], where they discuss the challenges of handling different types of diagrams, the complexity in representing the syntax and semantics, and handling noise. Noise in the graphic makes data extraction difficult, as the data points can be close and hence can be skewed or intersecting. Shahab et al. [6] presented the different techniques that were used to solve the problem of detecting and recognizing text from complex images. Relevant infor-

mation include understanding the axis labels, legend and the values the plots represent.

Attempts at graphics understanding from scientific plots can be seen in [1, 7] targeting bar charts, and simple line plots [3, 8, 9]. An understanding of the information required to be extracted is a key component in disambiguating the relevant section of a graphic, and [8] tackles the extraction of relevant information from line plots which are one of the more commonly used information graphics in a research setting. Additionally, in [3] the authors propose a method using hough transform and heuristics to identify the axis lines. The rules include the relative position between axis lines, the location of axis lines in a 2-D plot, and the relative length of axis lines in a 2-D plot. The textual information that is embedded into the graphics, such as axes labels and legends, in line plots and bar charts, is also crucial to understanding the graphics. Connected components and boundary/perimeter features [10] have been used to characterize document blocks. [11] discusses methods to extract numerical information in charts. They use a 3 step algorithm to detect text in an image using connected components, detect text lines using hough transform [12], and inter-character distance and character size to identify final text strings followed by Tesseract OCR [13] for text recognition. Color (HSV) based features have also been used to separate plot lines and text/axes [14] for interpretation of line plots that use color to discriminate lines. This study reports results on about 100 plots classified as ‘easy’ by the authors. They also use a color-based text pixel extraction scheme where the text is present only outside the axes and in the legend.

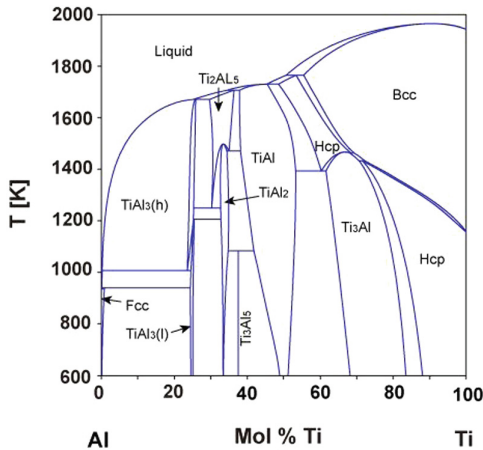


Fig. 2. Example of a challenging phase diagram.

Once the relevant data has been extracted, the next logical step is in connecting them in a coherent manner to interpret the information contained in the chart. [15] discusses the importance of communicative signals, which are information that can be used to interpret bar charts. Some of the signals of interest that

represent the information from the graphic include annotations, effort, words in caption, and highlighted components. Allen et. al [16], in one of the earliest works in the area developed a system for deducing the intended meaning of an indirect speech act. In [9], a similar idea is used in understanding line plots by breaking down each line plot into trends and representing each trend by a message. These constituent trend level messages are combined to obtain a holistic message for the line plot.

While phase diagrams belong broadly to the class of plots, they require special treatment due to the complex embedding of information into these diagrams, as explained in the next section. The contour nature of the plot, complex text placement with semantic import, and challenging locations and orientations coupled with the optional presence of other graphic symbols such as arrows that are vital for semantic interpretation of the figure, justify a dedicated exploration of such complex diagrams.

3 Phase Diagrams

Phase diagrams are graphs that are used to show the physical conditions (temperature, pressure, composition) at which thermodynamically distinct phases occur and coexist in materials of interest [17]. A common component in phase diagrams is lines which denote the physical conditions in which two or more phases coexist in equilibrium - these are known as *phase boundaries*. The X and Y axes of a phase diagram typically denote a physical quantity such as temperature, pressure and, in the case of alloys or mixtures, the ratio of components by weight or by molar fraction. As stated earlier, we focus on phase diagrams of binary metal alloys where the X-axis is molar fraction percentage and the Y axis is temperature. In Fig. 1, the blue lines within the plot denote the phase boundaries. All points bounded by a phase boundary represent physical conditions at which the material of interest, in this case an alloy of silver and zinc, occurs in the same phase. The name or label of this phase, for example α in Fig. 1, is typically present somewhere within the phase boundary. The various Greek letters present in the phase diagram represent different types of solid phases (i.e. crystal structures). Positioning within a phase defines the ratio of the different phases, as well as the composition of the phases. All of these characteristics heavily impact the material properties.

We can observe that there are several regions that are unlabeled. These regions represent multi-phase regions and the phases that constitute this mixture are obtained by using the phase labels of the regions to the left and right of the unlabeled region. Additionally, as shown in Fig. 2, in several phase diagrams, labels are sometimes provided to the phase boundary instead of the region. These cases represent intermetallic compositions (ie. the labeled phase exists only at that one composition, thus explaining the vertical line which is labeled). In such cases, the same rule to infer phase labels holds true except we would be using a vertical line on the left or on the right to obtain one of the two phase labels.

4 Overview of Our Approach

From a document analysis perspective, a phase diagram can be seen to consist mainly of alphanumeric text, often with accompanying Greek characters, in vertical and horizontal orientations; bounded regions of uniform phase within the plot; and descriptions of axes and numerical quantities along the axes. As can be seen in Fig. 2, narrow and small phase regions, presence of arrows, text located very close to phase boundaries and different orientations pose steep challenges to the automated analysis. The key steps in automated phase diagram analysis are listed below and will be elaborated in the sections that follow:

- detection and recognition of text used to label phases
- extraction of regions of uniform phases
- association of each phase region to appropriate labels
- detection and recognition of axes text in order to convert image coordinates to physical coordinates and detect elements of the binary alloy

5 Identifying Text and Phase Regions

The phase diagram images that we have considered in this study were obtained from a single source - the Computational Phase Diagram Database [18] from the National Institute of Materials Science, so that the phase labeling, and plot and image styling are consistent. We gathered about 80 different phase diagrams of binary alloys consisting of a number of common transition metals and main group metals. Each image was preprocessed by Otsu thresholding [19] and inverting it, so that background pixels are off and foreground pixels are on. We then extracted contours using the border following algorithm proposed by Suzuki and Abe [20]. The largest contour extracted corresponds to the box that defines the axes and the plot. The contours are then divided into plot (inside the largest contour), and non-plot (outside the largest contour). Plot and non-plot contours were manually annotated using an in-house annotation tool and a database of about 720 phase region contours and about 7100 text region contours were created. This database was divided into a training and validation set in the ratio 4 : 1. We then extracted the following features from the contours:

- Unit normalized coordinates of the contour bounding box
- Bounding box area normalized with respect to image area
- Contour area normalized with respect to area of image
- Convex hull area normalized with respect to area of image
- Ratio of contour area to convex hull area
- Ratio of contour area to bounding box area
- Ratio of convex area to bounding box area
- Contour perimeter normalized with respect to image perimeter
- Orientation, Eccentricity of contour
- Hu invariant moments

The features are designed so as to normalize the effect of large size disparity between phase regions and text regions. Orientation and eccentricity are computed using first and second order image moments of the contour. Hu invariant moments [21] are seven image moments that are invariant to rotation, translation, scale and skew. They are commonly used in object recognition and segmentation [22, 23].

Table 1. Confusion matrix for contour classification

%	Phase	Text
Phase	97.60	2.40
Text	0.15	99.85

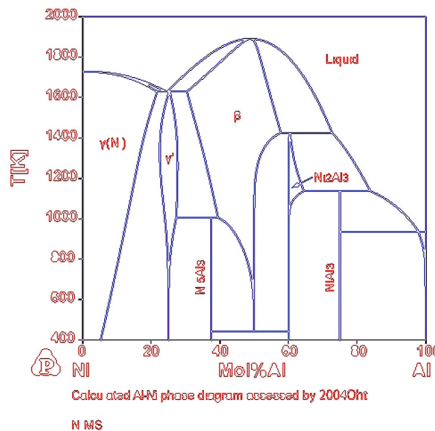


Fig. 3. Classification of text (red) and phase (blue) contours. Best viewed in color. (Color figure online)

The feature vector extracted from the contour has a dimensionality of 20. Features are extracted from training contours and a gradient boosted tree-based classifier is trained to classify between phase contours and text contours. We choose this classifier as we found it to be the most robust to our unbalanced data among other classifiers such as support vector machine (SVM), random forests and neural networks. The number of estimators was chosen to be 1000 and the maximum depth of each tree was fixed at 10 after a grid search. The performance of the trained classifier is evaluated on the validation set. Table 1 shows the confusion matrix obtained after evaluation. We can see that our model is quite proficient at classifying phase and text contours. Figure 3 shows the classification of text and phase contours in one of our phase diagram images. The text contours are marked in red and phase contours are marked blue.

6 Mapping Regions to Labels

After classification of all contours into non-phase and phase, we concentrate on grouping the text contours into words and recognizing the text, so that these word labels can then be mapped to the appropriate phase contours.

6.1 Segmenting Text into Words

As a first step, all text contours that are fully contained within another text contour are eliminated. Then, we sort all of the plot text contours in increasing order of y-coordinate of the centroid cy . A text contour i , whose centroid y-coordinate cy_i value is within a certain threshold H_t^1 from the previous contour $i - 1$ is grouped together as belonging to the same line. Otherwise, it becomes the start of a new line. Once the text contours are grouped into lines, we sort text contours in a single line by the increasing order of x-coordinates of their centroid cx . A text contour i is grouped together as belonging to the same word as the previous contour $i - 1$ if their centroid x-coordinates differ by a value of H_t^2 , otherwise it becomes the start of a new word. We fixed the values of H_t^1 as h_{mean} , the average height of all text contour bounding boxes within the plot and H_t^2 as $1.5 \times w_{mean}$, the average width of text contour bounding boxes in that particular line.

Using this method of line and word grouping works well for horizontally oriented text, however vertically aligned text are still left as isolated contours. In order to group vertical text, we repeat the procedure described above, except we group the text contours by x-coordinate of centroids to obtain vertical text lines and switch to grouping by y-coordinate of centroids to obtain vertical words in each line. We use a different set of thresholds, V_t^1 and V_t^2 respectively. We chose

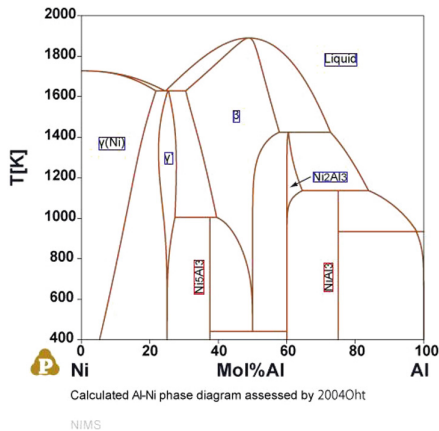


Fig. 4. Grouping of text into vertical (red) and horizontal (blue) words. Best viewed in color. (Color figure online)

the values of V_t^1 as w_{mean} , the average width of all isolated text contour bounding boxes and V_t^2 as $1.5 \times h_{mean}$, the average height of text contour bounding boxes in that particular vertical line.

Some single character text or contours that contain many characters may be left isolated and not grouped into horizontal or vertical lines. These contours are marked as ambiguous. The ambiguity is resolved by rotating the contours 90° in both clockwise and anti-clockwise directions and attempting to recognize the text in all three configurations. The configuration that yields the highest confidence is selected as the right orientation for these contours. Once the correct orientation of all contours is known, we perform OCR on all the words by using the orientation information. For vertically aligned text, we flip the word about the Y-axis, and compare recognition confidence in both directions to finalize the orientation. Figure 4 shows the grouping and orientation of plot text contours of the phase diagram in Fig. 3. Recognition is performed using the Tesseract library [13].

6.2 Detection of Arrows

Since the phase label for phase regions that are small in area cannot be placed within the region, they are usually displayed elsewhere and an arrow is used to indicate the region or line for the phase association. It is therefore necessary to identify arrows in order to accurately match these text contours to the corresponding phase contour. Arrows occur frequently in our dataset and are vital for correct interpretation of the phase diagram as can be seen in Fig. 2. We use a Hough line detector to detect arrows. Since the length of arrows varies in our dataset, we tune the Hough line detector to detect short line segments. Collinear and overlapping line segments are merged to yield the list of arrows in the image. The arrow direction is determined by comparing the center of mass of the arrow and its geometric midpoint. Due to more pixels located at the head of the arrow, we expect the center of mass to be between the head and the midpoint. For every arrow, we find the word region closest to the tail and the phase contour or vertical line closest to the head and these are stored as matched pairs. Figure 5 shows an example of successful arrow detection and corresponding text box association.

6.3 Completing the Mapping

Once the text within the plot is grouped and recognized and the arrows in the image have been dealt with, we proceed to associate the rest of the text labels to the appropriate phase regions and boundary lines. Vertical words are mapped to the nearest unlabeled vertical line by measuring the perpendicular distance between the centroid of the word bounding box and the line. After this, we match phase regions to horizontal text labels by finding the text bounding boxes that are fully contained within the phase region. We resolve conflicts, if any, by giving priority to text labels whose centroid is closest to that of the phase region. Labels for unlabeled phase contours are inferred using the rules described in Sect. 3.

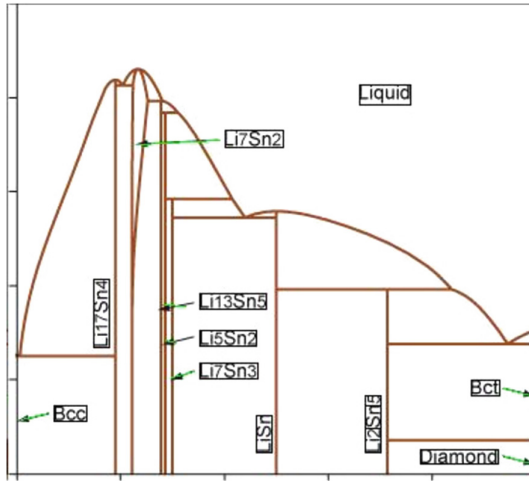


Fig. 5. Arrow detection and corresponding text box association. Best viewed in color. (Color figure online)

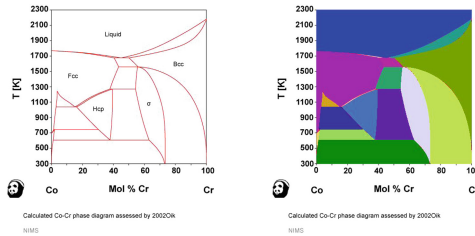


Fig. 6. Extraction of phase contours. Best viewed in color. (Color figure online)

6.4 Handling Text in the Axes

Text grouping, recognition and orientation determination is performed for text contours outside the plot boundary using the same procedure described in Sect. 6. The text regions to the left of the plot box closest to the top-left and bottom left corners are identified and recognized. Using the vertical distance between these two regions as well as the recognized numerals we can easily compute the value of the temperature that corresponds to the top-left and bottom-left corners of the plot box. Thus, we are able to translate the image coordinates (x, y) to the physical coordinates $(mole\ fraction, temperature)$. With this, we will be able to query any required physical coordinate for any binary alloy, convert it to image coordinates and find the phase contour which contains this point and return the label assigned to the contour.

7 Evaluation and Discussion

Despite designing a system which digitizes a phase diagram and returns the phase information of any queried point, we choose to eschew the traditional information retrieval oriented evaluation scheme. Instead, we present our accuracy of phase contour labeling for both cases - labels present within the phase diagram and labels that have to be inferred. This is because, in our case, we could potentially generate an infinite number of real-valued queries within the bounds of the plot axes and each one would have a corresponding phase label response. Depending on the kind of points queried we could have precision and recall numbers skewed to very high or very low accuracy and there would be no guarantee of a fair evaluation of our system. By measuring the accuracy of phase contour labeling, we can therefore obtain a comprehensive idea of the efficacy of our system.

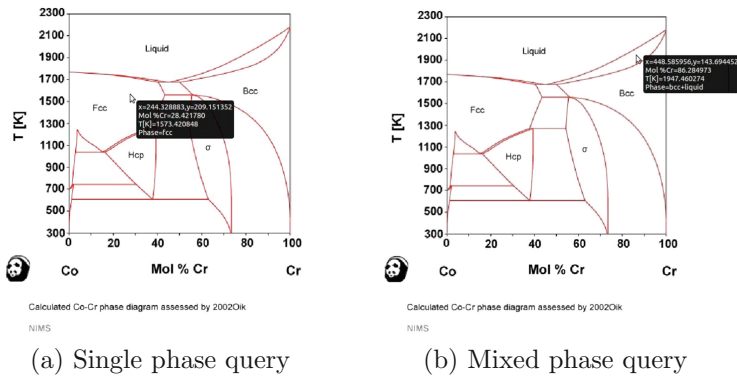


Fig. 7. Demo of our live phase query retrieval system. Best viewed in color. (Color figure online)

To this end, we have annotated the phase diagrams using the LabelMe annotation tool [24]. Expert annotators provided the text labels for all phase contours as well as relevant phase boundaries which were used as ground truth and the accuracy of the labels generated by our algorithm was measured against the truth. We report our accuracy of phase contour labeling for both cases - labels present within the phase diagram (94%) and labels that have to be inferred (88%).

We believe that the results show promise, as seen in Figs. 6 and 3. Our contour extraction and text classification works well even for varied contour sizes and shapes. A minimalistic demo application constructed using our methods is shown in Figs. 7(a) and (b), where we display the transformed physical coordinates as well as the phase of the material at the cursor position.

Discovery of Bulk Metallic-Glasses

Aside from the phase information, we also detect ‘eutectic points’ (see Sect. 1), which are point(s) in a phase diagram indicating the chemical composition and temperature corresponding to the lowest melting point of a mixture of components. These points serve as an important first order signature of alloy chemistries and are vital for design of ‘metallic-glasses’, a class of material of increasing interest and importance. The eutectic points can be determined by analyzing the smoothened contour of the liquid phase, for which both contour separation and accurate matching of label and region is critical. We also measure the so-called ‘eutectic angle’ corresponding to each eutectic point which is defined as the angle formed by the contour lines leading into and out of the eutectic point. An example is seen in Fig. 8. Blue circles are used to mark the location of eutectic points and the corresponding angles are shown nearby.

We analyzed a database of binary metallic phase diagrams and quantitatively defined that a deep eutectic angle is roughly between 0° and 75° . This value was defined by identifying the design rule which most correctly identified metallic glass forming compounds. This work therefore allows us to define the ‘deep eutectic’ in terms of a design rule, as opposed to the more general usage of the term to date. When combined with radii difference scaled by composition (the value along the X-axis) at the eutectic point, we were able to identify binary metallic systems that were likely to form metallic glasses. Following this analysis, we identified 6 different binary metallic systems to have a *high probability* of metallic-glass formation, which were previously unknown. A complete list of 38 different systems, with exact compositions and temperatures, which were previously unidentified as glass-forming, uncovered due to our work, are listed in [4].

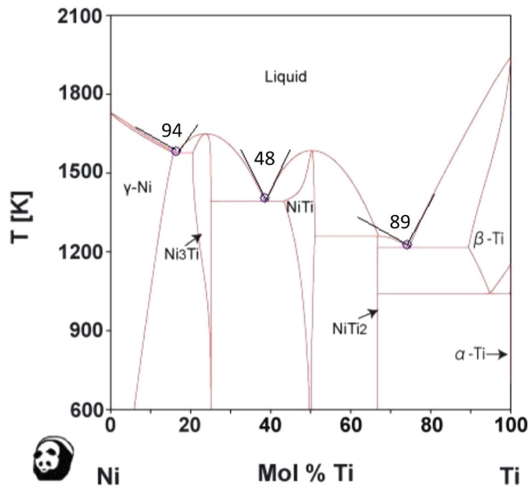


Fig. 8. Detection of eutectic points and angles in phase diagrams. Best viewed in color. (Color figure online)

Conclusion

Given the importance of a digitized phase diagram database to the materials community at large, we believe that our effort in developing automated tools to digitize phase diagrams from technical papers is a valuable contribution with significant impact. In the future, we would like to create a comprehensive, high-resolution database of phase diagrams, and improve label and phase region matching tasks. We would also like to extend our work to support the detection and storage of critical points and material parameters which are key in design and manufacture of certain materials. Further, the materials domain is rich in graphs, figures and tables that contain valuable information, which when combined and collated into large indexable, digital databases, would help the materials community to accelerate the discovery of new and exciting materials.

Acknowledgments. This material is based upon work supported by the National Science Foundation under Grant No.1640867 (OAC/DMR). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

1. Elzer, S., Carberry, S., Zukerman, I.: The automated understanding of simple bar charts. *Artif. Intell.* **175**(2), 526–555 (2011)
2. Carberry, S., Elzer, S., Demir, S.: Information graphics: an untapped resource for digital libraries. In: *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 581–588. ACM (2006)
3. Lu, X., Kataria, S., Brouwer, W.J., Wang, J.Z., Mitra, P., Giles, C.L.: Automated analysis of images in documents for intelligent document search. *IJDAR* **12**, 65–81 (2009)
4. Dasgupta, A., et al.: Probabilistic assessment of glass forming ability rules for metallic glasses aided by automated analysis of phase diagrams. *Scientific Reports* - under review (2018)
5. Blostein, D., Lank, E., Zanibbi, R.: Treatment of diagrams in document image analysis. In: Anderson, M., Cheng, P., Haarslev, V. (eds.) *Diagrams 2000*. LNCS (LNAI), vol. 1889, pp. 330–344. Springer, Heidelberg (2000). https://doi.org/10.1007/3-540-44590-0_29
6. Shahab, A., Shafait, F., Dengel, A.: ICDAR 2011 robust reading competition challenge 2: reading text in scene images. In: *International Conference on Document Analysis and Recognition* (2011)
7. Zhou, Y.P., Tan, C.L.: Bar charts recognition using hough based syntactic segmentation. In: Anderson, M., Cheng, P., Haarslev, V. (eds.) *Diagrams 2000*. LNCS (LNAI), vol. 1889, pp. 494–497. Springer, Heidelberg (2000). https://doi.org/10.1007/3-540-44590-0_45
8. Nair, R.R., Sankaran, N., Nwogu, I., Govindaraju, V.: Automated analysis of line plots in documents. In: *2015 13th International Conference on Document Analysis and Recognition (ICDAR)*, pp. 796–800. IEEE (2015)
9. Radhakrishnan Nair, R., Sankaran, N., Nwogu, I., Govindaraju, V.: Understanding line plots using bayesian network. In: *2016 12th IAPR Workshop on Document Analysis Systems (DAS)*, pp. 108–113. IEEE (2016)

10. Rege, P.P., Chandrakar, C.A.: Text-image separation in document images using boundary/perimeter detection. *ACEEE Int. J. Sig. Image Process.* **3**(1), 10–14 (2012)
11. Mishchenko, A., Vassilieva, N.: Chart image understanding and numerical data extraction. In: *Sixth International Conference on Digital Information Management (ICDIM)* (2011)
12. Duda, R.O., Hart, P.E.: Use of the hough transformation to detect lines and curves in pictures. *Commun. ACM* **15**(1), 11–15 (1972)
13. Smith, R.: An overview of the tesseract OCR engine. In: *ICDAR*, vol. 7, pp. 629–633 (2007)
14. Choudhury, P.S., Wang, S., Giles, L.: Automated data extraction from scholarly line graphs. In: *GREC* (2015)
15. Elzer, S., Carberry, S., Demir, S.: Communicative signals as the key to automated understanding of simple bar charts. In: Barker-Plummer, D., Cox, R., Swoboda, N. (eds.) *Diagrams 2006. LNCS (LNAI)*, vol. 4045, pp. 25–39. Springer, Heidelberg (2006). https://doi.org/10.1007/11783183_5
16. Perrault, C.R., Allen, J.F.: A plan-based analysis of indirect speech acts. *Comput. Linguist.* **6**(3–4), 167–182 (1980)
17. Campbell, F.C.: *Phase Diagrams: Understanding the Basics*. ASM International (2012)
18. Computational phase diagram database. cpddb.nims.go.jp/cpddb/periodic.htm. Accessed 06 Feb 2017
19. Otsu, N.: A threshold selection method from gray-level histograms. *Automatica* **11**(285–296), 23–27 (1975)
20. Suzuki, S., et al.: Topological structural analysis of digitized binary images by border following. *Comput. Vis. Graph. Image Process.* **30**(1), 32–46 (1985)
21. Hu, M.K.: Visual pattern recognition by moment invariants. *IRE Trans. Inf. Theory* **8**(2), 179–187 (1962)
22. Flusser, J., Suk, T.: Rotation moment invariants for recognition of symmetric objects. *IEEE Trans. Image Process.* **15**(12), 3784–3790 (2006)
23. Zhang, Y., Wang, S., Sun, P., Phillips, P.: Pathological brain detection based on wavelet entropy and hu moment invariants. *Bio-med. Mater. Eng.* **26**(s1), S1283–S1290 (2015)
24. Russell, B.C., Torralba, A., Murphy, K.P., Freeman, W.T.: LabelMe: a database and web-based tool for image annotation. *Int. J. Comput. Vis.* **77**(1), 157–173 (2008)