Learning to Generate Posters of Scientific Papers*

Yuting Qiang[1], Yanwei Fu[2], Yanwen Guo[1†], Zhi-Hua Zhou[1] and Leonid Sigal[2]
[1] National Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210023, China
[2] Disney Research Pittsburgh, 4720 Frobes Avenue, Lower Level, 15213, USA
{qiangyuting.new,ywguo.nju}@gmail.com, zhouzh@nju.edu.cn, {yanwei.fu,lsigal}@disneyresearch.com

Abstract

Researchers summarize and represent their paper content with scientific posters, which efficiently convey their ideas. Generating a good scientific poster, however, is challenging for novel researchers, since it needs to be *readable*, *informative*, and *aesthetic*. This paper for the first time studies the challenging problem of learning to generate posters from scientific papers. To this end, a data-driven framework is proposed by utilizing probabilistic graphical models. Specifically, given contents to display, the key elements of a good poster, including panel layout and attributes of each panel, are learned and inferred from data. Then composition of graphical elements within each panel is synthesized. To validate our framework, we contribute a Poster-Paper dataset with exhaustively labelled attributes of poster panels. Qualitative and quantitative results indicate the effectiveness of our framework.

Introduction

There emerge a large quantity of scientific papers in various academic conferences and journals every year. For example, CVPR 2016 accepted more than 600 papers. It is time-consuming to read all of these papers for a researcher in this area. Converting a paper into a poster representation is an important means to efficiently and compactly convey ideas, methods and information of the paper. It is thus essential to keep the poster readable, informative and aesthetic. However, it is challenging to design a high-quality scientific poster considering both utility and the aesthetic goal, especially for those researchers who are not proficient at design tasks. For convenience, they may just divide a poster into several columns then arrange contents sequentially. A poster designed by this way would look similar and dull to some extent. In general, poster design is a complicated and time-consuming task. Both understanding the paper contents and the experience in poster design are required.

Automatic tools for scientific poster generation would help researchers by providing them with an easier way to effectively share their research. Further, given the large amount of scientific papers on ArXiv and other on-line repositories, such tools may also provide a way for other researchers to consume the contents more easily. Rather than browsing the original papers, they may be able to browse the generated poster previews more efficiently.

However, in order to generate a valid scientific poster according to the original paper, many problems need to be solved: 1) Content extraction. Both important textual and graphical contents need to be extracted from the original paper; 2) Panel layout. Contents should fit the panel. Besides, the shape and position of each panel should be well arranged; 3) Graphical element (figure and table) arrangement. Within each panel, textual contents could be sequentially itemized. But for graphical elements, their sizes and positions should be carefully considered. Due to these challenges, there are few automatic tools for scientific poster generation.

In this paper, we propose a data-driven method to automatically generate a poster according to the paper content. Content extraction and layout generation are two key components to generate a satisfactory and visually pleasing poster. For content extraction, we use TextRank (Mihalcea and Tarau 2004) to extract textual contents, and we provide an user interface to extract graphical contents. Our approach focuses primarily on layout generation. We realize this in three steps. First, we propose a simple probabilistic graphical model to infer panel attributes. Second, we introduce a tree structure to represent the panel layout, based on which we further design a recursive algorithm to generate the panel layout. Third, in order to synthesize the layout within each panel, we train another probabilistic graphical model to infer the attributes of graphical elements.

Compared with posters designed by the authors, our approach can generate different results to adapt to different paper sizes by training our model with different dataset, thus expanding the richness and variety of layouts. To the best of our knowledge, this paper presents the first framework for poster generation from the original paper.

Our framework has the following contributions.

- Probabilistic graphical models are proposed to learn scientific poster design patterns, including panel attributes and graphical element attributes, from existing posters.

*This work is supported by NSFC (61333014, 61373059, and 61321491) and JiangsuSF (BK20150016).

†Corresponding author

- A new algorithm considering both information conveying and poster aesthetics is developed to generate the poster layout.
- We also contribute to the community a Poster-Paper dataset with exhaustively labelled attributes of posters.

Related Work

General Graphical Design. The general graphical design has been studied in computer graphics community. This involves several related, yet different topics, including *text-based layout generation* (Jacobs et al. 2003; Damera-Venkata, Bento, and O'Brien-Strain 2011; Hurst, Li, and Marriott 2009), *single-page graphical design* (O'Donovan, Agarwala, and Hertzmann 2014; Harrington et al. 2004), *photo album layout* (Geigel and Loui 2003), *furniture layout* (Merrell et al. 2011; Yu et al. 2011), and even *interface design* (Gajos and Weld 2005). Among them, text-based layout pays more attention on informativeness, while attractiveness also needs to be considered in poster generation. Other topics would take aesthetics as the highest priority. However, some principles (such as alignment or read-order) need to be followed in poster design. In summary, poster generation needs to consider readability, informativeness and aesthetics of the generated posters simultaneously.

Manga Layout Generation. Several techniques have been studied to facilitate layout generation for western comics or manga. For example, *scene frame extraction* (Arai and Herman 2010; Pang et al. 2014), *automatic stylistic manga layout generation* (Cao, Chan, and Lau 2012; Jing et al. 2015), and *graphical element composition* (Cao, Lau, and Chan 2014). For preview generation of comic episodes (Hoashi et al. 2011), both frame extraction and layout generation are considered. Other research areas, such as *manga retargeting* (Matsui, Yamasaki, and Aizawa 2011) and *manga-like rendering* (Qu et al. 2008) also draw considerable attention. However, none of these methods can be directly used to generate scientific posters, which is our focus in this paper.

Our panel layout generation is inspired by the recent work on manga layout (Cao, Chan, and Lau 2012). We also use a binary tree to represent the panel layout. By contrast, the manga layout trains a Dirichlet distribution to sample a splitting configuration, and different Dirichlet distribution for each kind of instance need to be trained. Instead, we propose a recursive algorithm to search for the best splitting configuration.

Overview

Problem Formulation. Assume that we have a set of posters $\mathbf{M}$ and their corresponding scientific papers. Each poster $m \in \mathbf{M}$ includes a set of panels $\mathbf{P}_m$. And each panel $p \in \mathbf{P}_m$ has a set of graphical elements (figures and tables) $\mathbf{G}_p$. Each panel p is characterized by five attributes:

text length (l_p) text length of a panel;

text ratio (t_p) text length of a panel relative to text length of the whole poster, $t_p = l_p / \sum_{q \in \mathbf{P}_m} l_q$;

graphical elements ratio (g_p) [1] the size of graphical elements within a panel relative to the total size of graphical elements in the poster;

panel size (s_p) **and aspect ratio** (r_p) $s_p = w_p \times h_p$ and $r_p = w_p / h_p$, where w_p and h_p denote the width and height of a panel with respect to the whole poster, separately.

Each graphical element $g \in \mathbf{G}_p$ has four attributes:

graphical element size (s_g) $s_g = w_g \times h_g$;

graphical element aspect ratio (r_g) $r_g = w_g / h_g$ where w_g and h_g denote the width and height of a graphical element relative to the whole paper respectively;

horizontal position (h_g) we assume that the panel contents are arranged sequentially from top to down[2]. Thus only relative horizontal position is considered here which is defined as an enumerated variable chosen from the set $\{left, center, right\}$;

graphical element size in poster (u_g) the ratio of the width of one graphical element to the panel width.

To learn how to generate the poster, our goal is to *determine the above attributes of each panel* p *and each graphical element* $g \in \mathbf{G}_p$*, as well as to infer the arrangement of all panels.*

Intuitively, a trivial solution is to use a learning model (e.g., SVR) to learn how to regress these attributes, including s_p, r_p, u_g, and h_g, while regarding t_p, g_p , l_p, r_g, and s_g as features. However, such a solution lacks an insight mechanism for exploring the relationships between the panel attributes (e.g., s_p) and graphical element attributes (e.g., u_g). And it may fail to meet the requirements of readability, informativeness, and aesthetics. We thus propose a novel framework to solve our problem.

Overview. To generate a *readable*, *informative* and *aesthetic* poster, we simulate the rule-of-thumb on how people design the posters in practice. We generate the panel layout, then arrange the textual and graphical elements within each panel.

Our framework overall has four steps (as shown in Figure 1) while the core of our framework focuses on three specific algorithms designed to facilitate poster generation. We first extract textual contents from the paper using TextRank (Mihalcea and Tarau 2004)[3], this will be detailed in the Experimental Result section; while non-textual contents are extracted by user interaction. All these extracted contents are sequentially arranged and represented as the first blob in Figure 1. Initial inference of the panel attributes (i.e. panel size s_p and aspect ratio r_p) is then conducted by learning a probabilistic graphical model from the training data. Furthermore, *panel layout* is synthesized by developing a recursive algorithm to further update these attributes and generate

[1] Note that there is a little difference between this variable and text ratio t_p. We do not use the figure size in poster. Instead, we use the corresponding figure from the original paper.

[2] This is true when using latex beamer to make posters.

[3] We use TextRank for context extraction, which nevertheless can be replaced with other state-of-the-art textual summary algorithms.

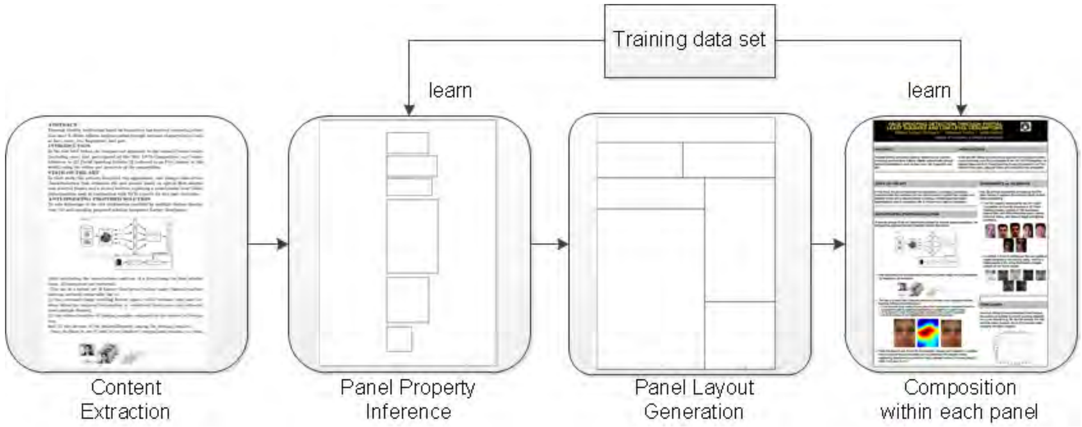


Figure 1: Overview of the proposed approach.

an *informative* and *aesthetic* panel layout. We finally compose panels by utilizing the graphical model to further synthesize the visual properties of each panel (such as the size and position of those graphical elements).

Methodology

Panel Attributes Inference. Our approach tries to divide a scientific poster into several blocks and each of them is a rectangular panel. Each panel should not only be with an appropriate size to contain textual and graphical contents, but also be in a suitable shape (aspect ratio) as far as aesthetics is concerned. Our approach learns a probabilistic graphical model to infer the initial values of the size and aspect ratio of one panel.

As each panel is composed of both textual description and graphical elements, we can safely assume that panel size (s_p) and aspect ratio(r_p) rely on text ratio t_p and graphical element ratio g_p. Therefore, the likelihood of a set of panels p is

$$Pr(s_p, r_p|t_p, g_p) = \prod_{p \in P} Pr(s_p|t_p, g_p) Pr(r_p|t_p, g_p) \quad (1)$$

where $Pr(s_p|t_p, g_p)$ and $Pr(r_p|t_p, g_p)$ are conditional probability distributions (CPDs) of s_p and r_p, given t_p and g_p. We define them as two conditional linear Gaussian distributions:

$$Pr(s_p|t_p, g_p) = N(s_p; \mathbf{w_s} \cdot [t_p, g_p, 1]^\mathsf{T}, \sigma_\mathbf{s}) \quad (2)$$

$$Pr(r_p|t_p, g_p) = N(r_p; \mathbf{w_r} \cdot [t_p, g_p, 1]^\mathsf{T}, \sigma_\mathbf{r}) \quad (3)$$

where t_p and g_p can be decided in the *content extraction* step demonstrated in Figure 1. $\mathbf{w_s}$ and $\mathbf{w_r}$ are the parameters that leverage the influence of various factors; $\sigma_\mathbf{s}$ and $\sigma_\mathbf{r}$ are the variances of models. These parameters ($\mathbf{w_s}$, $\mathbf{w_r}$, $\sigma_\mathbf{s}$ and $\sigma_\mathbf{r}$) are estimated by maximum likelihood from training data. We can then infer the initial attributes of each panel directly.

Note that in order to learn from limited data, this step actually employs two assumptions: (1) s_p and r_p are conditionally independent; (2) The attribute set of each panel is independent of others. Nevertheless, we need the panels to be neither too small in size (s_p) nor too distorted in aspect ratio (r_p), to ensure readable, informative and aesthetic poster generated. Thus, these two assumptions are introduced here. Furthermore, the attribute value set estimated from this step aims at giving a good initial evaluation of the property of each panel. We use the next two steps to further relax these assumptions and discuss the relationship between s_p and r_p, as well as the connections of different panels (Algorithm 1).

To ease exposition, we denote the set of panels as $L = \{(s_{p_1}, r_{p_2}), (s_{p_2}, r_{p_2}), \cdots, (s_{p_k}, r_{p_k})\}$, where s_{p_i} and r_{p_i} are the size and aspect ratio of ith panel p_i, separately; with $|L| = k$.

Panel Layout Generation. One conventional way to design posters is to simply arrange them in two or three columns style. This scheme, although simple, however, makes all posters look similar and unattractive. Inspired by manga layout generation (Cao, Chan, and Lau 2012), we propose a more vivid panel layout generation method. Specifically, we arrange the panels with a binary tree structure to help represent the panel layout. Thus, the panel layout generation is formulated as a process of recursively splitting of one page, as illustrated and explained in Figure 2.

Conveying information is the most important goal for a scientific poster, thus we maintain the relative size for each panel during panel layout generation. As for the aesthetic goal, we define the loss function for the panel shape variation,

$$l(p_i) = |r_{p_i} - r'_{p_i}| \quad (4)$$

where r'_{p_i} is the aspect ratio of a panel after optimization. This will lead to an aesthetic loss function for poster-level panel shape variation,

$$Loss(L, L') = \sum_{i=1}^{k} l(p_i) \quad (5)$$

where L' is the poster panel set after optimization. In the splitting step, the combinational choices of splitting positions

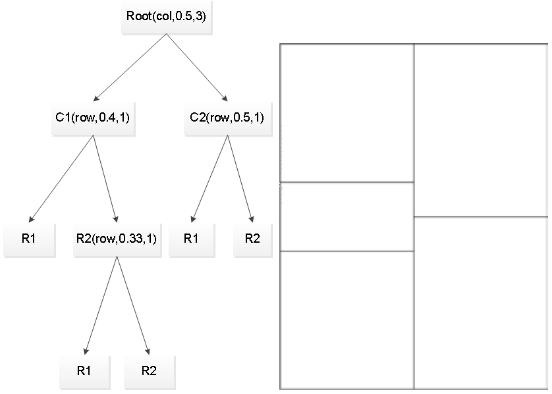


Figure 2: Panel layout and the corresponding tree structure. The tree structure of a poster layout contains five panels. The first splitting is vertical with the splitting ratio (0.5, 0.5). The poster is further divided into three panels in the left, and two panels in the right. This makes the whole page as two equal columns. For the left column, we resort to a horizontal splitting with the splitting ratio (0.4, 0.6). The larger one is further horizontally divided into two panels with the splitting ratio (0.33, 0.67). We only split the right column once, with the splitting ratio (0.5, 0.5).

can be recursively computed and compared on the loss function above. We choose the one with the lowest loss (Eq. 5). The whole algorithm is summarized in Algorithm 1.

Composition within a Panel. Now we can generate the composition of a panel which fixes the attributes of graphical elements of the poster. To infer these attributes, we still employ the probabilistic graphical model. Particularly, the key attributes we need to estimate are the horizontal position h_g and graphical element size u_g. In our model, horizontal position h_g relies on s_p, l_p and s_g, while u_g relies on r_p, s_g and r_g, so the likelihood for this graphical model is

$$Pr(h_g, u_g | s_p, r_p, l_p, s_g, r_g) = \prod_{p \in P} \prod_{g \in p} Pr(u_g | s_p, l_p, s_g) Pr(h_g | r_p, s_g, r_g) \quad (6)$$

$Pr(u_g | s_p, l_p, s_g)$ and $Pr(h_g | r_p, s_g, r_g)$ are the conditional probability distributions (CPDs) of u_g and h_g given s_p, l_p, r_p, s_g and r_g respectively. The conditional linear Gaussian distribution is also used,

$$Pr(u_g | s_p, l_p, s_g) = N(u_g | \mathbf{w_u} \cdot [s_p, l_p, s_g, 1]^\mathsf{T}, \sigma_\mathbf{u}) \quad (7)$$

where $\mathbf{w_u}$ is the parameter to balance the influence of different factors. Since we take horizontal position h_g as an enumerated variable, a natural way to estimate it is to make it as a classification problem by using the softmax function,

$$Pr(h_g = i | r_p, s_g, r_g) = \frac{e^{\mathbf{w_h}_i \cdot [r_p, s_g, r_g, 1]^\mathsf{T}}}{\sum_{j=1}^{H} e^{\mathbf{w_h}_j \cdot [r_p, s_g, r_g, 1]^\mathsf{T}}} \quad (8)$$

where H is the cardinality of the value set of h_g, i.e. $H = 3$. $\mathbf{w_h}_i$ is the ith row of $\mathbf{w_h}$. The maximum likelihood method

---

Algorithm 1 Panel layout generation

Input:
Panels which we learned from graphical model
$L = \{(s_{p_1}, r_{p_1}), (s_{p_2}, r_{p_2}), \cdots, (s_{p_k}, r_{p_k})\}$;
rectangular page area x, y, w, h.
Output:

```
if k == 1 then
    adjust panels[0] to adapt to the whole rectangular page area, return the aesthetic loss: |r_{p_0} − w/h|;
else
    for each i ∈ [1, k − 1] do
        t = Σ_{j=1}^{i} s_{p_j} / Σ_{j=1}^{n} s_{p_j};
        Loss_1 = Panel Arrangement((s_{p_1}, r_{p_1}), (s_{p_2}, r_{p_2}), ··· , (s_{p_i}, r_{p_i}), x, y, w, h × t);
        Loss_2 = Panel Arrangement((s_{p_{i+1}}, r_{p_{i+1}}), (s_{p_{i+2}}, r_{p_{i+2}}), ··· , (s_{p_k}, r_{p_k}), x, y + h × t, w, h × (1 − t));
        if Loss > Loss_1 + Loss_2 then
            Loss = Loss_1 + Loss_2;
            record this arrangement;
        end if
        Loss_1 = Panel Arrangement((s_{p_1}, r_{p_1}), (s_{p_2}, r_{p_2}), ··· , (s_{p_i}, r_{p_i}), x, y, w × t, h);
        Loss_2 = Panel Arrangement((s_{p_{i+1}}, r_{p_{i+1}}), (s_{p_{i+2}}, r_{p_{i+2}}), ··· , (s_{p_k}, r_{p_k}), x + w ∗ t, y, w × (1 − t), h);
        if Loss > Loss_1 + Loss_2 then
            Loss = Loss_1 + Loss_2;
            record this arrangement;
        end if
    end for
end if
return Loss and arrangement.
```

---

is used to estimate these parameters, which include $w_\mathbf{u}$,$\mathbf{w_h}$ and $\sigma_\mathbf{u}$.

Different from Eq. (1), directly inferring h_g and u_g is not advisable, since the panel content may exceed the panel bounding box and affect the aesthetic measure of a poster. To avoid this problem, we employ the likelihood-weighted sampling method (Fung and Chang 1990) to generate samples from the model, by maximizing the likelihood function (Eq. 1 and Eq. 6) with this strict constraint,

$$\sum_{g \in p} h_p \times u_g + \alpha \times \beta \times l_p / w_p < h_p \quad (9)$$

where α and β denote the width and height of a single character respectively. The first term of the above constraint indicates the height of graphical elements while the second term represents the height of textual contents.

Experimental Results

Experimental Setup. We contribute the first Poster-Paper dataset to the community on this problem. Specifically, we have selected 25 well-designed pairs of scientific papers and their corresponding posters from 600 public available papers and posters. These papers are all about scientific topics,

stage		Average time
Text extraction⋆		28.81s
Panel attributes inference	learn	0.85s
	infer	0.01s
Panel layout generation		0.13s
Composition within panel	learn	2.17s
	infer	0.03s+19.09s⋆

Table 1: Running time of each step. ⋆: it takes 0.03s for inference computation and then 19.09s for latex file generation.

and their posters have relatively similar design styles of panels. We further annotate some key attributes, such as panel width, panel height and so on. We make a training and testing split: 20 for training and five for testing. There is total of 173 panels in our dataset. 143 for training and 30 for testing.

We use TextRank to extract textual contents from the original paper, in order to weight the importance values of different sections, we can set different extracting ratio for each section. Thus according to the informativeness rule, an important section is usually advised to use more contents, thus given a bigger panel. For simplicity, this paper uses equal important weights. And user-interaction is also involved to highlight and select the important figures and tables from the original paper. We use the Bayesian Network Toolbox (BNT) (Murphy 2002) to estimate some key parameters. To infer graphical element attributes, we generate 1000 samples by the likelihood-weighted sampling method (Fung and Chang 1990) for Eq. 6 while the constraint Eq.9 is used. With the inferred meta data, the final poster is generated in latex Beamerposter format with Lankton theme.

For baseline comparison, we invite three second-year Phd students, who are not familiar with our project, to design posters for the test set. These three students work on computer vision and machine learning and have not yet published any papers on these topics and thus are novices to research. Given the testing papers, we ask the students to work together and design a poster for each paper.

Running Time. Our framework is very efficient. Our experiments were done on a PC with an Intel Xeon 2.0 GHz CPU and 144GB RAM, Tab. 1 shows the average time we needed for each step. Strictly speaking, we can not compare with "previous methods", since we are the first work on poster generation and there is no existing work to compare. Nevertheless, we argue that the total running time is significantly less than that people usually take when they design a good poster, and it is also less than the time spent to generate a poster by three novices in the following Quantitative evaluation section.

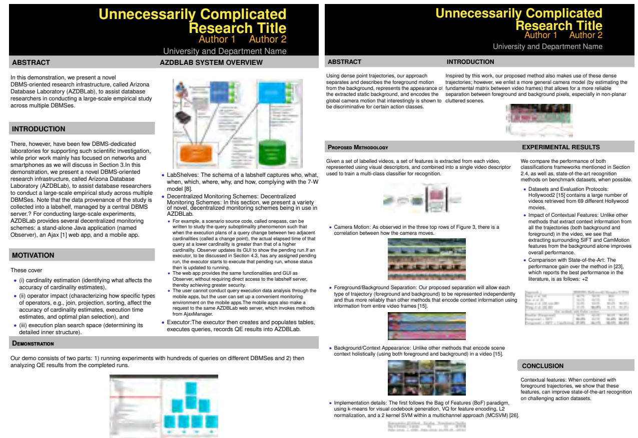
Quantitative Evaluation. We quantitatively evaluate the effectiveness of our approach.

(1) Effectiveness of panel inference. For this step, we compare the inferred size and aspect ratio of panels with the trivial solution – SVR which trains a linear regressor[4] to predict the panel size and panel aspect ratio from training data. We use the panel attributes from the original posters[5] as the ground-truth and compute the mean-square error (MSE) of inferred values versus ground-truth values. Our results can achieve 3650.4 and 0.67 for panel size and aspect ratio. By contrast, the values of SVR method are 3831.3 and 0.76 respectively. This shows that our algorithm can better estimates the panel attributes than SVR.

(2) User study. User study is employed to compare our results with the original posters and posters made by novices. We invited 10 researchers (who are experts on the evaluated topic and kept unknown to our projects) to evaluate these results on readability, informativeness and aesthetics. Each researcher is sequentially shown the three results generated (yet with different order) and is asked to score their judgements from $0-10$, where 0, 5 and 10 indicate the lowest, middle and highest scores of corresponding metrics. The final results are averaged for each metric item.

As shown in Tab. 2, our method is comparable to the original posters on *readability* and *informativeness*; and it is significantly better than posters made by novices. This validates the effectiveness of our method, since the inferred panel attributes and generated panel layout will save most valuable and important information. By contrast, our method is lower than the original posters on aesthetics metric (yet, still higher than those from novices). This is reasonable, since aesthetics is a relatively subjective metric and it generally needs to involve much human interaction and adjustment. It is an open problem to generate more aesthetic posters from papers.

Qualitative Evaluation of Three Methods. We qualitatively compare our result (Figure 3(b)) with the poster by a novice in Figure 3(a) and the original poster Figure 3(c). All

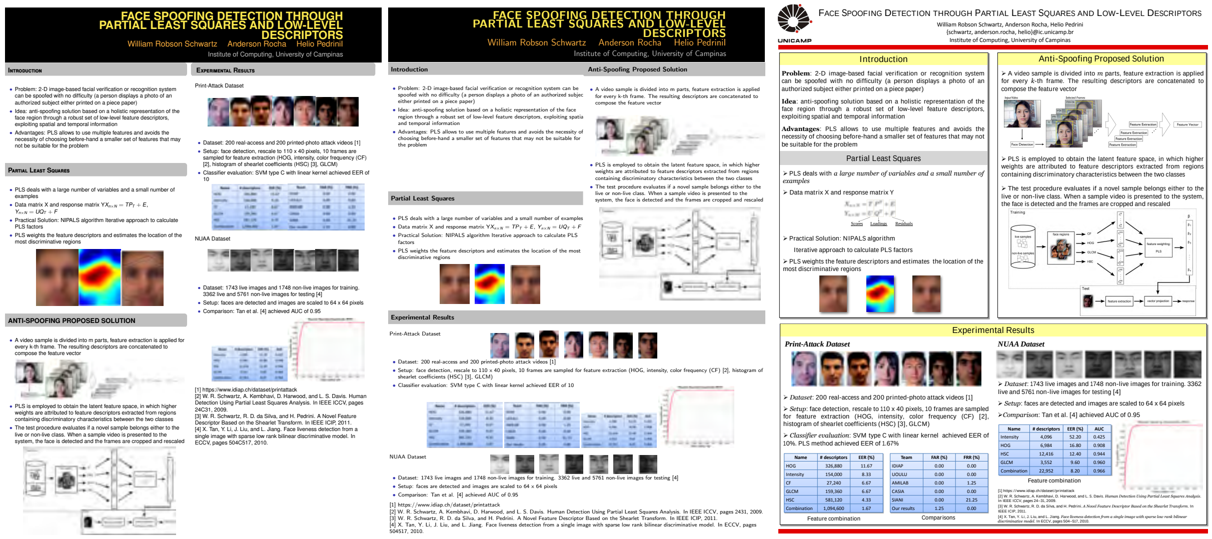


(a) (b)

Figure 4: Qualitative comparison of our result (b) and novice's result (a). Please refer to our supplemental material for larger size figures.

[4] s_p and r_p are used as features for SVR. The parameters are cross-validated. Nonlinear kernels (such as RBF) make worse results due to the over-fitting of training data.

[5] Note that though the panels of original poster may not be the best ones, they are the best candidates to serve as the ground-truth here.



(a) Designed by novice (b) Our result (c) Original poster

Figure 3: Results generated by different ways

Metric	Readability	Informativeness	Aesthetics	Avg.
Our method	6.94	7.06	6.86	6.95
Posters by novices	6.69	6.83	6.12	6.54
Original posters	7.08	7.03	7.43	7.18

Table 2: User study of different posters generated.

of them are for the same paper.

It is interesting to show that if compared with the panel layout of original poster, our panel layout looks more similar to the original one than the layout by novices. This is due to, first, the Poster-Paper dataset has a relative similar graphical design with high quality, and second, our split and panel layout algorithms work well to simulate the way how people design posters.

Further Qualitative Evaluation. We further qualitatively evaluate our results (Figure 4) by the general graphical design principles (O'Donovan, Agarwala, and Hertzmann 2014), i.e., *flow*, *alignment* and *overlap and boundaries*.

Flow. It is essential for a scientific poster to present information in a clear read-order, i.e. readability. People always read a scientific poster from left to right and from top to bottom. Since Algorithm 1 recursively splits the page of a poster into *left* and *right* or *top* and *bottom*, the panel layout we generate ensures that the read-order matches the section order of the original paper. Within each panel, our algorithm also sequentially organizes contents which also follow the section order of original paper and this improves the readability.

Alignment. Compared with the complex alignment constraint in (O'Donovan, Agarwala, and Hertzmann 2014), our formulation greatly simplifies it by using an enumerate variable to indicate the horizontal position of graphical elements h_g, namely *left*, *center* and *right*. This simplification does not spoil our results which still have reasonable alignment as illustrated in Figure 4 and quantitatively evaluated by three different metrics in Tab. 2.

Overlap and boundaries. Overlapped panels will make the poster less readable and less aesthetic. To avoid overlapping, our approach (1) recursively splits the page for panel layout; (2) sequentially arranges the panels; (3) enforces the constraint Eq. 9 to penalize the cases of overlapping between graphical elements and panel boundaries. Thus our algorithm can achieve reasonable results without significant overlapping and crossing boundaries. Similar to those manually aligned poster – Figure 3(c), our results (Figure 3(b)) are free from significantly overlapped panels and boundaries.

Conclusion and Future Work

Automatic tools for scientific poster generation are important for poster designers. They will save a lot of time with this kind of tools. Design is a hard work, especially for scientific poster design, which should take into account both utility and the aesthetic goal. Many principles about scientific poster design can not help design work directly. By contrast, we propose an approach to learn design pattern from existing examples, and this approach will hopefully lead to an automatic tool for scientific poster generation to aid designers.

Except for scientific poster design, our approach also provides a framework to learn other kinds of design patterns, for example web-page design, single-page graphical design, and so on. And by providing different set of training data, our approach could generate different layout layout style. Our work has several limitations. We do not consider font types in our current implementation and only adopt a simple yet effective aesthetic metric. We plan to address these problems in future.

Acknowledgements

We would like to thank the anonymous reviewers for their insightful suggestions in improving this paper.

References

Arai, K., and Herman, T. 2010. Method for automatic e-comic scene frame extraction for reading comic on mobile devices. In *Information Technology: New Generations (ITNG), 2010 Seventh International Conference on*, 370–375. IEEE.

Cao, Y.; Chan, A. B.; and Lau, R. W. H. 2012. Automatic stylistic manga layout. *ACM Trans. Graph.* 31(6):141:1–141:10.

Cao, Y.; Lau, R. W.; and Chan, A. B. 2014. Look over here: Attention-directing composition of manga elements. *ACM Transactions on Graphics (TOG)* 33(4):94.

Damera-Venkata, N.; Bento, J.; and O'Brien-Strain, E. 2011. Probabilistic document model for automated document composition. In *Proceedings of the 11th ACM symposium on Document engineering*, 3–12. ACM.

Fung, R. M., and Chang, K.-C. 1990. Weighing and integrating evidence for stochastic simulation in bayesian networks. 209–220.

Gajos, K., and Weld, D. S. 2005. Preference elicitation for interface optimization. In *Proceedings of the 18th annual ACM symposium on User interface software and technology*, 173–182. ACM.

Geigel, J., and Loui, A. 2003. Using genetic algorithms for album page layouts. *IEEE multimedia* (4):16–27.

Harrington, S. J.; Naveda, J. F.; Jones, R. P.; Roetling, P.; and Thakkar, N. 2004. Aesthetic measures for automated document layout. In *Proceedings of the 2004 ACM symposium on Document engineering*, 109–111. ACM.

Hoashi, K.; Ono, C.; Ishii, D.; and Watanabe, H. 2011. Automatic preview generation of comic episodes for digitized comic search. In *Proceedings of the 19th ACM international conference on Multimedia*, 1489–1492. ACM.

Hurst, N.; Li, W.; and Marriott, K. 2009. Review of automatic document formatting. In *Proceedings of the 9th ACM symposium on Document engineering*, 99–108. ACM.

Jacobs, C.; Li, W.; Schrier, E.; Bargeron, D.; and Salesin, D. 2003. Adaptive grid-based document layout. 22(3):838–847.

Jing, G.; Hu, Y.; Guo, Y.; Yu, Y.; and Wang, W. 2015. Content-aware video2comics with manga-style layout. *Multimedia, IEEE Transactions on* 17(12):2122–2133.

Matsui, Y.; Yamasaki, T.; and Aizawa, K. 2011. Interactive manga retargeting. In *ACM SIGGRAPH 2011 Posters*, 35. ACM.

Merrell, P.; Schkufza, E.; Li, Z.; Agrawala, M.; and Koltun, V. 2011. Interactive furniture layout using interior design guidelines. *ACM Transactions on Graphics (TOG)* 30(4):87.

Mihalcea, R., and Tarau, P. 2004. Textrank: Bringing order into texts. Association for Computational Linguistics.

Murphy, K. 2002. Bayes net toolbox for matlab.

O'Donovan, P.; Agarwala, A.; and Hertzmann, A. 2014. Learning layouts for single-pagegraphic designs. *Visualization and Computer Graphics, IEEE Transactions on* 20(8):1200–1213.

Pang, X.; Cao, Y.; Lau, R. W.; and Chan, A. B. 2014. A robust panel extraction method for manga. In *Proceedings of the ACM International Conference on Multimedia*, ACM MM.

Qu, Y.; Pang, W.-M.; Wong, T.-T.; and Heng, P.-A. 2008. Richness-preserving manga screening. 27(5):155.

Yu, L.-F.; Yeung, S.-K.; Tang, C.-K.; Terzopoulos, D.; Chan, T. F.; and Osher, S. J. 2011. Make it home: automatic optimization of furniture arrangement. *ACM Transactions on Graphics (TOG)-Proceedings of ACM SIGGRAPH 2011, v. 30, no. 4, July 2011, article no. 86.*