SSII2024 30周年記念特別企画 技術マップ: レイアウト生成 / クリエイティブビジョン

by 北田, 長内, 土井, 吉橋, ツウンニア (LINEヤフー株式会社)

技術マップ　文献情報

レイアウト生成

新聞・雑誌・論文 etc. のグリッドベースのデザインの起こり (1920 ~ 1940)

グリッドベースのシステムの発展 [Hurlburt, 1977, 1978; Muller-Brockman 1981]

"グリッドをどのようにレイアウトしていくか?" という問いから研究が発展

AdaGridDoc [Jacobs+ TOG'03]

DesignLayout [O'Donovan+ TVCG'14]

Label Layout Styles [Vollick+ UIST'07]

DesignSpace [O'Donovan+ CHI'15]

Furniture Layout Guidelines [Merrell+ SIGGRAPH'11]

レイアウトの網羅的な評価尺度の提示

レイアウト生成初期の取り組み

LayoutVAE [Jyothi+ ICCV'19]

NDN [Lee+ ECCV'20]

CanvasVAE [Yamaguchi ICCV'21] — Crello データセットの公開

Coarse-to-Fine [Jiang+ AAAI'22]

Text2Poster [Jin+ ICASSP'22]

READ [Patil+ CVPR-W'19]

GANベース ➜ VAEベースへの注目の高まり

ContentGAN [Zheng+ SIGGRAPH'19] — Magazine データセットの公開 ポスター生成と密接に関係

LayoutGAN [Li+ ICLR'19]

AC-LayoutGAN [Li+ TVCG'21]

NNベースの生成手法の発展と評価指標の定着

Transformer encoder を用いた BERT-like モデル

LayoutGCN [Manandhar+ ECCV'20]

レイアウト生成に対するグラフ畳み込みモデルの応用

RUITE [Rahman+ IUI'21]

CLG-LO [Kikuchi+ ACMMM'21]

BLT [Kong+ ECCV'22]

LayoutTransformer [Gupta+ ICCV'21]

VTN [Arroyo+ CVPR'21]

LayoutFormer++ [Jiang+ CVPR'23]

LayoutDETR [Yu+ arXiv'22] — Ads Banner データセットの公開

Parse-Then-Place [Lin+ ICCV'23] — WebUI データセットの公開

Transformer Decoder を用いた AR ベースモデル

画像品質の評価指標をレイアウト生成にも応用

FID [Heusel+ NeurIPS'17] → Fidelity & Diversity [Naeem+ ECCV'20]

LLM に HTML/JSON 形式でレイアウトを出力させる etc.

ChatGPT 公開 [OpenAI 2022/11] → LayoutGPT [Feng+ NeurIPS'23] → LayoutNUWA [Tang+ ICLR'24]

Llama2 公開 [Meta 2023/07]

生成するレイアウト数の依存を解消するように拡張

DLT [Levi+ ICCV'23]

LayoutDM [Inoue+ CVPR'23]

LayoutDiffusion [Zheng+ CVPR'23]

Dolfin [Wang+ arXiv'23]

PLay [Cheng+ ICML'23]

LDGM [Hui+ CVPR'23]

LayoutDiffusion [Zhang+ ICCV'23]

LayoutDM [Chai+ CVPR'23]

LACE [Chen+ ICLR'24]

CoLay [Cheng+ arXiv'24]

VLC [Shabani+ CVPR'24]

連続 / 離散の定式化による拡散モデルの提案

LayoutPGGAN [Ishiyama+ IJCOM'23]

Learn & Sample [Wang+ IJCAI'23]

LayoutDM/LayoutDiffusion を冠する提案が2023/03 ~ 05 に集中

レイアウト生成・ポスター生成の今後の課題

- データセットの大規模化・多様化
- 人間の感性に近い評価指標の構築
- レイアウト生成・ポスター生成における基盤モデルとなりうるモデルの構築
- 言語モデルの構造に囚われない自然なマルチモーダルモデルへの拡張

提案手法をすぐさま改善 ポスター生成へ拡張

ポスター生成

レイアウト生成

N 個の視覚的要素 (図、表、文章 etc.) の集合を元にキャンバス上へ適切に配置するタスク

条件なし生成

- ノイズからレイアウトを生成
- 集合の要素を適当に並べて自己回帰的に生成
- 要素をランダムにマスクしその要素を当てる

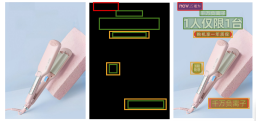
条件あり生成

- 要素のカテゴリ
- 要素のカテゴリとサイズ
- 要素の関係性
- レイアウトのリファインメント
- 指定の要素群を元に要素を補完

ポスター生成

コンテンツを考慮した条件ありレイアウト生成

- 顕著性検出等を用いて画像中の顕著な部分を取得
- 顕著領域を考慮してレイアウト生成を学習・推論



PosterLayout [Hsu+ CVPR'23]

- 古典手法 (not NN)
- GAN (Generative Adversarial Network) ベース
- VAE (Variational Auto-Encoder) ベース
- GNN (Graph Neural Network) ベース
- AR (Auto-regressive) ベース
- Diffusion ベース
- LLM (Large Language Model) ベース
- Flow ベース

ICVT [Cao+ ACMMM'22]

CreGAN [Wang+ ACMMM'22]

PosterLayout データセットの公開

CGL データセットの公開

CGL-GAN [Zhou+ IJCAI'22]

ポスター生成用データセットの公開 ➜ 分野の発展に貢献

PosterLayout [Hsu+ CVPR'23]

FlexDM [Inoue+ CVPR'23]

マスク言語モデルの学習

Two-stage LayoutDM [Chai+ ACMMM'23]

RALF [Horita+ CVPR'24]

Desigen [Weng+ arXiv'24]

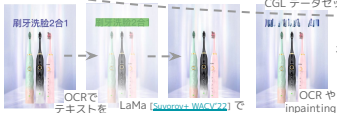
LayoutFlow [Guerreiro+ arXiv'24]

CGL データセットのテキストを拡張

RADM [Li+ CIKM'23]

COLE [Jia+ arXiv'23]

Graphist [Cheng+ arXiv'24]



本 PosterLayout の図の例は [Horita+ CVPR'24] より

LayoutPrompter [Li+ NeurIPS'23]

PosterLlama [Seol+ CVPR-W'24]

PosterLLaVa [Yang+ arXiv'24]

広告画像作成支援の枠組み ➜ ポスター生成の設定と同様

Vinci [Guo+ CHI'21]

De-rendering Stylized Texts [Shimoda+ ICCV'21]

SmartText [Li+ TMM'22] — 自然画像上にテキストを配置する試み

ポスター生成のシステム構築

AutoPoster [Lin+ ACMMM'23]

TextPainter [Gao+ ACMMM'23]

Planning & Rendering [Li+ arXiv'23]

PostDoc [Jaisankar+ arXiv'24]

サーベイ

Design Principles for Visual Communication [Agrawala+ Comm. of ACM'11]

A Survey for Graphic Design Intelligence [Hung+ arXiv'23]

Understanding Design Collaboration: A Systematic Literature Review [Shi+ CSCW'23]

Intelligent Layout Generation based on Deep Generative Models [Shi+ Information Fusion '23]

~2000　2010　2020　2021　2022　2023　2024　年