

**PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance**

This is to certify that the thesis/dissertation prepared

By Ali Jahanian

Entitled
Quantifying Aesthetics of Visual Design Applied to Automatic Design

For the degree of Doctor of Philosophy

Is approved by the final examining committee:

JAN P. ALLEBACH

RUTH ROSENHOLTZ

MIREILLE BOUTIN

S.V.N. VISHWANTATHAN

M J T . SMITH

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification/Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

JAN P. ALLEBACH

Approved by Major Professor(s): _____

Approved by: M. R. Melloch

12/08/2014

Head of the Department Graduate Program

Date

QUANTIFYING AESTHETICS OF VISUAL DESIGN
APPLIED TO AUTOMATIC DESIGN

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Ali Jahanian

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

December 2014

Purdue University

West Lafayette, Indiana

To my mother and father for their unlimited love and support, for believing in me, and for teaching me the value of hard work, honesty, and creativity.

To my brothers Mohammad and Mostafa for their unlimited love and support.

To my lovely sister Mina.

ACKNOWLEDGMENTS

I would like to sincerely thank Professor Jan P. Allebach for believing in me, facilitating to nurture my creativity, and investing on me to develop as a researcher. Throughout my studies I learned many wise lessons from him and his monitoring discipline and I hope one day to pay it back to the academic society.

I would also like to sincerely thank Professor S.V.N. Vishwanathan for his support and lessons in statistics. Throughout my collaboration with him, I learned many wise lessons about research, scientific work ethics, and perspective on life.

I would like to acknowledge Professor M.J.T. Smith for his valuable guidance and kind support.

I would like to acknowledge Professor Mimi (Mireille Boutin) for the lessons that I learned in her pattern recognizer course, and Dr. Ruth Rosenholtz for her thoughtful comments and exposing me to a different point of view in visual cognition.

I would also like to acknowledge the faculty members of School of Design at Purdue University, especially Professor David Sigman and Professor Petronio Bendito for their valuable insights to my work and the fact that they always motivated me to follow my passion in quantifying aesthetics and automatic design.

Finally, I would like to acknowledge Mrs. Cornelia Haeringen as one of my best teachers for proofreading this thesis, which was accompanied by kindness and motivation. I would also like to acknowledge Ms. Elana Cutter for kindly helping me with collecting design datasets as well as editing some parts of my publications.

TABLE OF CONTENTS

	Page
LIST OF TABLES	ix
LIST OF FIGURES	x
ABSTRACT	xv
1 INTRODUCTION	1
1.1 Motivation	1
1.2 Thesis Contributions	2
1.2.1 Identifying Challenges	2
1.2.2 Defining Research Strategies	3
1.2.3 Establishing A Taxonomy	4
1.2.4 Compiling Design Guidelines	5
1.2.5 Automating Design Processes	5
1.2.6 Semantic Design Mining	5
1.2.7 Quantitatively Revisiting Two Design Theories via Large-scale Data	6
1.2.8 Deploying Crowdsourcing for Design	6
1.2.9 Devising Recommendation Systems for Design	7
2 ON THE LEGITIMACY OF QUANTIFYING AESTHETICS	8
2.1 Theoretical Considerations	8
2.2 Taxonomy	10
2.2.1 Automatic Approaches	11
2.2.2 Human Inspection Approaches	13
2.3 Conclusion	14
3 DESIGN MINING COLOR SEMANTICS	15
3.1 Introduction	15
3.2 Theory	18

	Page
3.2.1 Color Cognition	18
3.2.2 Color Naming	19
3.2.3 Color Meanings and Semantics	19
3.2.4 Color Semantics, Emotions, and Preferences	20
3.2.5 Color Semantics and Cross-cultural Considerations	21
3.2.6 Color Semantics in Applied Arts	21
3.2.7 Color Semantics in HCI	22
3.3 Related Work	23
3.3.1 Color Semantics and Meanings	23
3.3.2 LDA Topic Modeling	24
3.3.3 User Studies via Crowdsourcing	24
3.3.4 Click Modeling	25
3.4 Data Collection	25
3.4.1 Preprocessing	26
3.4.2 Word Vocabulary	26
3.5 Statistical Model	27
3.5.1 Review of LDA-dual Model for Color Semantics	29
3.5.2 Inference	32
3.5.3 Effect of Color Basis	35
3.6 Interpreting Model Output	35
3.6.1 From Color Histograms to Color Palettes	36
3.6.2 From Weighted Bag of Words to Word Clouds	37
3.7 User Study	38
3.7.1 Formative Study	40
3.7.2 Summative Study	41
3.8 Interpreting the User Study	42
3.8.1 Statistical Model	42
3.8.2 Analyzing the Results	46

	Page
3.8.3 Association Directionality	47
3.9 Applications	49
3.9.1 Color Palettes Selection Using Semantics	49
3.9.2 Design Example Recommendation	50
3.9.3 Image Retrieval Using Color Semantics	51
3.9.4 Image Color Selection Using Semantics	51
3.9.5 Image Recoloring Using Semantics	53
3.10 Conclusion and Future Work	53
3.10.1 Visual Design Language and User Interaction	55
3.10.2 Quantifying Aesthetics	55
4 DESIGN MINING VISUAL BALANCE	56
4.1 Introduction	56
4.2 Theory	58
4.2.1 Visual Balance in Spatial Composition	58
4.2.2 Visual Rightness	58
4.3 Modeling Framework	60
4.3.1 EM for Gaussian Mixtures	63
4.3.2 Dataset	63
4.4 Results	63
4.5 Discussion and Future Work	65
5 AUTOMATIC DESIGN OF SELF-PUBLISHED MEDIA: A CASE STUDY OF MAGAZINE COVERS	67
5.1 Introduction	67
5.2 Theoretical Considerations	69
5.2.1 Principles of Design	69
5.2.2 Elements of Design	71
5.3 Related Work	73
5.3.1 Automated Layout	74

	Page
5.3.2 Color Design	74
5.3.3 Typography	75
5.3.4 Computational Aesthetics	75
5.4 Automatic Design System	76
5.4.1 Software Framework Overview	76
5.5 Layout	78
5.5.1 Defining Space	79
5.5.2 Visual Balance Considerations	80
5.6 Typography	83
5.6.1 Masthead	84
5.6.2 Headline and Byline	85
5.6.3 Cover Lines	86
5.7 Design of Color	88
5.7.1 Masthead Color	90
5.7.2 Cover lines Colors	92
5.7.3 Color Aesthetics and Text Legibility	93
5.8 Experimental Results	98
5.8.1 Results	99
5.9 Conclusion and Future Work	101
6 RECOMMENDATION SYSTEM FOR AUTOMATIC DESIGN	103
6.1 Introduction	103
6.2 Scenario	105
6.3 Related Work	106
6.4 Software Framework	106
6.4.1 Schematic View	107
6.5 Input User Interface	108
6.6 Evaluation of Input Photos	109
6.7 Designs User Interface	113

	Page
6.8 Personalization of Designs	115
6.9 Experimental Results	116
6.10 Conclusion and Future Work	117
REFERENCES	121
A Appendix Design Mining Color Semantics	140
B Appendix Derivations of Gibbs Sampling for LDA-dual	149
B.1 Notation	149
B.2 Joint Probability	150
B.3 Integrating Out Multinomials	151

LIST OF TABLES

Table	Page
A.1 Summary of our magazine covers dataset. In total 2,645 magazine covers from 71 different titles were collected. These titles are roughly assigned to 12 different genres of magazines by us for the purpose of our analysis.	140
A.2 Stop word list. The list of words that we exclude from the words on the covers to. Note that these words are inspected by manually visiting the first 30 words in the word topics inferred by the model.	143

LIST OF FIGURES

Figure	Page
1.1 Thesis research strategies: a schematic view.	4
2.1 Taxonomy of quantifying aesthetics: A schematic view. Refer to [19–28]. . .	11
3.1 Overview of our end-to-end framework for design mining of color semantics.	16
3.2 Pink is used in all of these designs, despite the fact that each of these designs belongs to a different context and class of magazines. Images from [167–169].	28
3.3 Hierarchical Bayesian plate model for the LDA-dual model, which combines color and word topics. D is the number of magazine covers; K is the number of color-word topics; N_d and M_d are the number of color swatches and words, respectively, in the d -th magazine cover.	30
3.4 LDA is both a generative and inference model. This image is inspired by [172].	31
3.5 Color-word topics inferred by the LDA-dual model. Illustration of the 12 color topics in the middle, their corresponding 12 word topics, 6 on top for the first 6 color histograms from the left, and the other 6 at the bottom. Note that for visualization, only the principle elements in the histograms are shown. Also note that the numerical weights of the word topics are shown next to the each word topic histogram. Similarly, the numerical weights of the words in each word topic histogram are shown next to their corresponding words. See Fig. 3.6 for the proportions of the 12 color-word topics for the magazine covers in the dataset.	33
3.6 Topics vs. titles. The proportion of each of the 12 color-word topics, k_1 to k_{12} (see Fig. 3.5) for each magazine title is illustrated. Note that the colors here are just legends for the purpose of visualization, and not related to the color-word topics.	34
3.7 Visualization process for the inferred color-word topics. To visualize the color-word topics histograms inferred by the model (see Fig. 3.5), we use 5-color palettes and word clouds as proxies to color histograms and word histograms, respectively.	36
3.8 Three first closest color palettes to each color histogram, in the pool of color palettes. Eq 6.3 describes the used similarity metric.	38

Figure	Page
3.9 The 12 color-word topics produced by the model are visualized in this figure. The top panel shows the color histogram, the second and third color panels show the top two color palettes we extracted from this histogram. The word topics are visualized in the bottom panel as word clouds, with the size of a word being proportional to its weight.	39
3.10 Flow of the summative user study.	42
3.11 Relevance matrices \hat{R}^1 and \hat{R}^2 for the first and second set of questions, respectively. The elements of these matrices are estimated intrinsic relevance of associations between colors and words, calculated from the participants' responses. The higher the value, the greater the intrinsic relevance associated by the users. Ideally, the diagonals should contain the highest values.	45
3.12 Results of examining the words to colors association directionality. Relevance matrices \hat{R}^1 and \hat{R}^2 for the first and second set of questions, respectively. The elements of these matrices are estimated intrinsic relevance of associations between words (rows) and colors (columns), calculated from the participants' responses. Compare with matrices in Fig. 3.11.	48
3.13 Application of color semantics in color palettes selection, and design example retrieval. See Applications section. Magazine cover images from [168, 187].	50
3.14 Image color selection using color semantics. a) The original image, b) colors that contribute to "travel" and "shop", and c) "women" and "fashion", in the original image. Image from the homepage of tripadvisor.com [192].	52
3.15 Pattern colorization using color semantics. a) The original pattern, b) "science", c) "shop", and d) "sports" and "men" colorized patterns. Images colorized using techniques in [194]. Pattern from <i>ColourLovers ArrayOfLilly</i> [148].	53
4.1 Arnheim's structural net. This net illustrates the hotspots on any visual artwork. Image reproduced from [44].	59
4.2 An example for overlap of saliency map with Arnheim's net.	60
4.3 Results of our GMM fitting.	64
5.1 Elements of a magazine cover. (image from [273])	71
5.2 A schematic view of ADoMC. Image reproduced from [110].	76
5.3 A general template for magazine covers.	79

Figure	Page	
5.4	Layout decisions and visual balance considerations. a) An image. b) GBVS [164] saliency map. c) Binarized mask from (b). d) A general template for magazine covers. e) Fusion of (c) and (d). ADoMC keeps the green region and discards the brown region, since the green region looks empty and is a good candidate to hold the cover lines. This decision is made by computing the visual balance as shown in (f). f) Computation of visual balance: different weights are assigned to the cells of the yellow grids on the left and right sides of the cover image: the darker the cells, the heavier the weights. Summing up the product of each cell's weight and the number of salient pixels (here in white) gives us the weights of the left and right grids (see Eq. 5.2). The side with a lighter weight will be considered as an empty space to hold more elements, such as the texts of the cover lines (see Eq. 5.3). Photos courtesy of Amir Garachorlou.	81
5.5	Examples of automatic cover lines text insertion on a magazine cover. a) The green region is a candidate to hold cover lines. We want to insert the text of the cover lines in such a fashion that the text indentation follows the shape of the green region and the size of the text is as large as possible. b) Result of our automatic text insertion in the green region in (a). c) Another result of our automatic text insertion in the green region in (a). Note that the number of words and different words in the cover lines yield different results.	87
5.6	Insertion of cover lines on a magazine cover. The blue boxes represent the bounding box of each line of the text and the orange geometric regions represent the candidate regions for inclusion of the text. a) A smaller size of the text may result in a higher fidelity to the candidate region of the text whereas b) shows that larger size of text is also desired. c) shows the distance penalties (in red and violet) and other variables. These distance penalties represented by red and violet arrows need to be minimized.	88
5.7	Elements of color design. a) Itten's hue wheel. Note that complementary colors are in opposite directions: for instance, yellow and violet, or orange and blue. b) Matsuda's harmonious templates. This figure is from [41]. Hues which are in shaded sectors are considered to be in harmony. The angles of the sectors are described in [40] and [41].	89
5.8	Our choice of hue wheel for computing the complementary color. a) A continuous pure hue wheel. This means that saturation and value are equal to the highest possible amounts. b) The inner wheel shows the hue wheel. But we discard the sectors marked with "×". As a result, we use the outer wheel to return the computed complementary (see Eq. 5.10). c) Our design of color palette will have a triple of hues, all in the shaded sectors of V-type and Y-type templates of Matsuda.	91

Figure	Page
5.9 Matsudas harmonious tone templates. a) Tone templates. This figure is from [41]. Tones which are in the shaded sectors are considered to be in harmony. The measures of the sectors are described in [41].	94
5.10 Text legibility solution and results using hue modification. a) Flowchart representing color modifications for text when the lightness (in $CIEL^*a^*b^*$) of the text and its local background are too close. In this flowchart, the lightness of the masthead is denoted by L_M , and the mean lightness of the masthead's local background by $L_M^{\overline{BG}}$. Also, HL and BL indicate headline and byline, respectively. The numerical thresholds are based on psychophysical experiments, some described in [285]. b) and c) Before and after applying the flowchart in (a), respectively. Photos courtesy of Andy Jessop.	96
5.11 Text legibility solution using lightness modification. a) Flowchart representing color modifications for text when the hue (in $CIEL^*a^*b^*$) of the text and its local background are too close. In this flowchart, the hue of masthead is denoted by H_M , and the mean hue of the masthead's local background by $H_M^{\overline{BG}}$. Also, HH and BH indicate headline and byline, respectively. The numerical thresholds are based on psychophysical experiments, some described in [285]. b) and c) Before and after applying the flowchart in (a). Photos courtesy of Andy Jessop.	97
5.12 Text legibility solution using shadow effect. An alternative solution to improve text legibility by applying shadow effects to the text, a) before, and b) after modifications, respectively. Photos courtesy of Andy Jessop.	98
5.13 A screenshot of the graphical user interface of ADoMC (Automatic Design of Magazine Covers) system.	99
5.14 Automatic designs generated by ADoMC. Photos courtesy of Sean Molin (a, b, and c), and Martin O'Connell (d).	100
6.1 A schematic view of R-ADoMC. The user uploads his/her photos, cover lines, and masthead. Then, he/she chooses some recommended descriptors to indicate the desired color mood of his/her design. R-ADoMC evaluates the user's photos and chooses the best. It then sends the selected photos to the design engine. The generated designs will be shown to the user, who can rate the generated designs. The ratings will be considered for the user's future designs to recommend him/her one more design based on his/her style.	107
6.2 Descriptors for color mood of design. The user can tell R-ADoMC which color mood he/she wishes to convey in his/her design. These descriptors are suggested by [43] to relate the world of colors to the world of styles and human preferences.	108

Figure	Page
6.3 Kobayashi’s Color Image Scale. Image taken from [43]. The contours show 15 patterns, labeled by all-capital letters. Note that Kobayashi has defined 1170 3-color combinations on this scale, however for simplicity only some of them are shown. We use these patterns to describe the desired color mood of the design.	110
6.4 The location of an image on Kobayashi’s Color Image Scale. This image is considered to be CHIC and MODERN according to this scale. The scale is taken from [43].	114
6.5 Designs UI for visualization of the designs generated in R-ADoMC. A rating option is available for each design, and the ratings will be considered in the user’s future designs. The “your style” design is a representation of the idea of personalized designs for the user. The more that the user works with the system, the more that the system is able to personalize the future designs. . .	115
6.6 Illustrations of magazine covers that were automatically designed on Kobayashi’s Color Image Scale (the scale image is taken from [43]). The cover images of these designs have been selected according to Kobayashi’s Patterns to represent the mood of the magazine cover designs. Images reproduced from [138]. Photos courtesy of Andy Jessop.	117
A.1 Relevance matrices. (a) and (b) females, (c) and (d) males, (e) and (f) Non-US participants, (g) and (h) US participants, (i) and (j) Non-designers, (k) and (l) designers.	144
A.2 Formative study. For this study one 5-color palette and the 12 word triples (under the color histograms) along with the “None of the above” option were shown. The participants were asked to choose as many as word triples that apply for the given 5-color palette. For each color histogram the first three closest 5-color palettes are shown in order, from top to bottom. Also for each color histogram a word triple as the representative of each word topic is shown.	145
A.3 Demography statistics of the summative user study. a) 14 highest participants’ countries. b) Participants’ age. c) Participants’ education.	146
A.4 Color topics histograms for $K = 24$	147
A.5 Color topics histograms when the color basis is the 3-color combinations designed by Kobayashi [43]. Note that the black bandings are the result of small proportions. That is, when one zooms in on the visualization software tool, he/she can see these small proportions of the 3-color combinations.	148

ABSTRACT

Jahanian, Ali PhD, Purdue University, December 2014. Quantifying Aesthetics of Visual Design Applied to Automatic Design. Major Professor: Jan P. Allebach.

In today's "Instagram" world, with advances in ubiquitous computing and access to social networks, digital media is adopted by art and culture. In this dissertation, we study what makes a good design by investigating mechanisms to bring aesthetics of design from realm of subjection to objection. These mechanisms are a combination of three main approaches: learning theories and principles of design by collaborating with professional designers, mathematically and statistically modeling good designs from large scale datasets, and crowdsourcing to model perceived aesthetics of designs from general public responses. We then apply the knowledge gained in automatic design creation tools to help non-designers in self-publishing, and designers in inspiration and creativity. Arguably, unlike visual arts where the main goals may be abstract, visual design is conceptualized and created to convey a message and communicate with audiences. Therefore, we develop a semantic design mining framework to automatically link the design elements, layout, color, typography, and photos to linguistic concepts. The inferred semantics are applied to a design expert system to leverage user interactions in order to create personalized designs via recommendation algorithms based on the user's preferences.

1. INTRODUCTION

In this chapter, I first motivate why it is important to quantify aesthetics of visual design and why there is an urge for automatic design. I then propose a collection of critical challenges that I confronted through my research. I finally conclude this chapter with my thesis research strategies and contributions.

1.1 Motivation

Visual design of self-publishing media over the internet is becoming an essential part of today's communication. As Bodker [1] argues, we are witnessing the so called third wave of Human Computer Interaction —HCI— in the information era, where conceptually and theoretically, aesthetics, as well as cognitive and emotional experience are an inevitable part of cultural levels. With the advances in ubiquitous computing and access to computers and social networks from anywhere at anytime, from home and via mobile devices, digital media has been adopted by art and culture in our everyday life [2].

In contrast to the first and second wave, design is no longer exclusively created on professional machines in enterprises, nor is its exhibition restricted to prints, galleries, or museums. There is an increasing request for software applications that provide users with easy-to-follow steps for creating visual media, such as photo albums, scrapbooks, magazines, blogs, and other forms of self-publishing media.

Despite the amount of visual communication and the huge corpora of already designed media on the internet, unfortunately, good design is still not accessible to the general public. Good design cannot be learned in an instant, and this presents a barrier for individuals and small to medium businesses who want to create good designs but cannot afford to hire a professional designer.

The gap between designers and non-designers is becoming more apparent. Many companies are now offering services for semiautomatic/automatic design. These services, however, are mainly for content delivery rather than being faithful to designing aesthetically pleasing media with adopted principles of design. This resonates that “bad design is everywhere, good design is invisible;” —a well-known concept among designers.

1.2 Thesis Contributions

In order to tackle the topic of my thesis, I identified three fundamental challenges. I present a set of questions to further define each of these challenges. I then describe the research strategies that I adopted to address these challenges. Finally, I enumerate the contributions of this thesis.

1.2.1 Identifying Challenges

(I) How to Quantify a Good Design?

- Is it legitimate to quantify aesthetics? Aesthetics seems to be subjective and design is so diverse.
- How to evaluate if a design is good: Is it aesthetically appealing? Is it successful in solving a visual problem? How to quantify and measure aesthetics of a design?

(II) How to Formulate Automatic Visual Design?

- How to understand a design process, from ideating, conceptualizing, to creating a design by means of design foundations, layout, color, and typography?
- How to recommend alternatives of design for a specific visual problem? How to understand the user’s need and purpose of his/her design? Can we recommend a personalized design to the user?

(III) How to Use Knowledge Gained to Improve Design for Masses?

- Who are the users of automatic design? Can we help non-designers to design better? Can we help designers to make better decisions by inspiring them with examples of good designs? Can we use automatic design for inspiration and creativity?
- How to use this knowledge gained to improve our algorithms, and define more meaningful, universal, and local sets of interactions in the design of user interfaces?
- How to develop a design language using aesthetics, principles of design, and crowd feedback to engage designers and non-designers, scientists, and other experts in solving a visual problem? How to engage both designers and clients in a design process?

1.2.2 Defining Research Strategies

Because this thesis deals with visual design and quantifying it and we aim to develop an autonomous system for design, I have approached these challenges in an interdisciplinary fashion. Figure 1.1 illustrates a schematic view of my research strategies. My research has been shaped by collaborating with professional designers. I ran interviews with several magazine art directors, and attended some design courses in the School of Design at Purdue University. In part, I strived to investigate well-known theories that professionals refer to and apply to their work. The third component is a set of principles in design inferred from the work of designers. I have applied these principles in order to develop an automatic design system. Moreover, I customised these principles to define high level features for both semantic design mining from work of professionals and for recommending design alternatives to the user.

As a specific case study of design, I focused on design of magazine covers. Designers consider magazine cover design as one of the richest kinds of design where all of the design elements, space, photo, logo, type, and color coalesce to convey a unified message. I

refer to our automatic design of magazine covers system as *ADoMC*. I then developed a recommendation system on top of *ADoMC* which I refer to as *R-ADoMC*. The main goal is to investigate the realization of a recommendation system for design, by identifying and addressing the main challenges in understanding the user's purpose of design, learning from his/her preferences, and predicting his/her style of design. Similarly, I refer to our data-driven system as *D-ADoMC*. This system is still under investigation. We will publish our work on how to incorporate our design mining findings with *ADoMC* in future work.

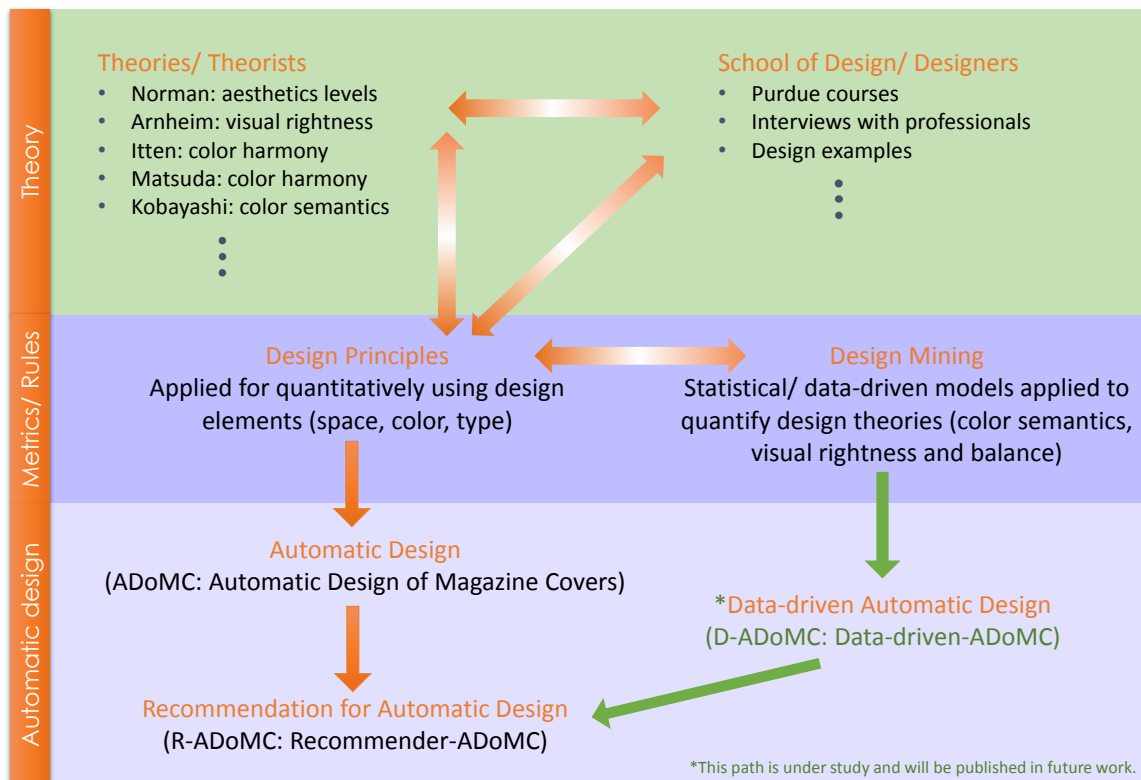


Fig. 1.1.: Thesis research strategies: a schematic view.

1.2.3 Establishing A Taxonomy

In my studies I found there are several different professions investigating aesthetics: mathematicians who strive to model aesthetics, engineers who aim to automate design, psychologists and neuroscientists who conduct qualitative experiments on aesthetics, and

designers who move aesthetics forward and occasionally propose theories on aesthetics. I strived to discover and explore the seminal work of these professions and established a taxonomy to present similarities and diversities of their work. See Chapter 2 for this taxonomy and theoretical considerations of quantifying aesthetics.

1.2.4 Compiling Design Guidelines

In order to quantify aesthetics and develop autonomous design systems, we need design guidelines. General design principles and guidelines along with well-known design theories may be found in books and are taught in schools of design. I learned from many books, but I also collaborated with professionals to interpret, customize, and apply some design guidelines for the purpose of this thesis. The lessons that I learned in design courses and in interviewing professional designers, art directors, journalists, and people at the faculty of School of Design at Purdue University are collected and presented throughout this thesis, and more specifically in Chapter 5.

1.2.5 Automating Design Processes

In this thesis, I present a framework for automatic design of magazine covers which I refer to as ADoMC. This framework is a consolidation of algorithmically quantifying the main modules in design: layout and using space, color design, and typography (see Chapter 5). I then deploy ADoMC as the core engine for our recommender ADoMC which I refer to as R-ADoMC (see Chapter 6).

1.2.6 Semantic Design Mining

Following advances in data mining and the notion of big data, the notion of design mining is emerging. In this thesis, I not only describe two frameworks for design mining of colors and visual balance, but more importantly, I propose semantic design mining. Arguably, visual communication design has to be both appealing and convey a message. This

elucidates that semantics and communicating with linguistic concepts by means of colors and other design elements should be taken into account. In Chapter 3, I describe a semantic design mining framework which specifically models semantics as a combination of colors and linguistic concepts. This framework can be applied for visual balance semantics, and typography semantics as well. Furthermore, this framework can leverage the idea of developing a visual language for design. Nevertheless, visual design language is not the main focus of this thesis and hence we consider it to be a part of our future work.

1.2.7 Quantitatively Revisiting Two Design Theories via Large-scale Data

There are a number of well-know theories in design. One example is the notion of color semantics. In Chapter 3, I revisit this notion by means of large-scale datasets and feedback of crowd. Another example is the theory of visual rightness and balance. This theory has been the subject of many qualitative research experiments. In Chapter 4, I describe how we quantitatively validate this concept by investigating a large-scale dataset of images. I believe that evaluating prior design theories via large-scale datasets, under the condition that these datasets are collected from the work of professionals, can scientifically contribute to the field of design.

1.2.8 Deploying Crowdsourcing for Design

Although design cannot be learned in an instant and we aim to learn from professionals by quantifying aesthetics of their work, arguably, even non-designers can distinguish if a design is a bad design. That is, we can learn from the wisdom of the crowd and collect responses about aesthetics of design. In particular, I have deployed crowdsourcing to evaluate the results of our color semantics modeling framework (see Chapter 3).

1.2.9 Devising Recommendation Systems for Design

Along with the advances in machine learning techniques and mechanisms for gathering users' personalized data, we are witnessing the emergence of recommendation systems in different areas such as marketing and education. Such a recommendation system for design is proposed in this thesis in Chapter 6. I refer to this system as *R-ADoMC*. A design recommendation system should be able to recommend to the user a number of designs and to customize the future designs for a user based on his/her personal preferences. Therefore, designing effective sets of interactions for user participation in the process of automatic design of visual media based on his/her preferences is required. More importantly, a recommendation design system has to map user's responses to design decisions. In other words, the set of features that should be defined are very different from other recommendation systems: we need to define features that are meaningful in design, for instance, the relation between colors. In Chapter 6, I use especially the notion of color emotions in a recommender system for automatic design of magazine covers, *R-ADoMC*, to recommend to the user a design based on his/her color emotion preferences. I also deploy the notion of color semantics for recommendation of design assets such as color palettes and design example retrieval in Chapter 3.

2. ON THE LEGITIMACY OF QUANTIFYING AESTHETICS

Is it legitimate to quantify aesthetics? Aesthetics is thought to be a matter of subjectiveness, on the other hand design may not be found in formulas. For instance, Jason K. McDonald states “One reason design formulas and routines are inadequate is because design so often addresses the untried, the unproven, and the unknown” [3]. However, we believe that with advances in computing, the phenomenon of big data, and more importantly, the fundamental difference between art and visual design, it is possible to quantify aesthetics of visual design.

In this chapter, we first discuss the theoretical considerations about aesthetics, and then offer a taxonomy of quantifying aesthetics in the fields of computer science/ engineering, psychology, and neuroscience.

2.1 Theoretical Considerations

Aesthetics of art is historically a subject of dispute between philosophers and art theorists. Korsmeyer in her book *Aesthetics: the big questions* [4] collects some of the fundamental questions about aesthetics of art, such as what is aesthetics, can we learn from aesthetics and art, who decides what is art and what is aesthetically pleasing. Korsmeyer also collects several essays and different viewpoints of philosophers to elucidate thinking about the aesthetics questions.

Korsmeyer summarizes two opposite doctrines about aesthetics, one which conceptualizes aesthetics as a matter of pleasure and taste; and the other which considers aesthetics as a matter of cognition and hence learning processes. Korsmeyer accounts for the former by referring to earlier philosophers such as Plato in his book the *Republic* and later philosophers such as Immanuel Kant and his *Critique of Judgement*; and for the latter, by referring to Aristotle in his *Poetics* and contemporary philosophers like Hans-Georg

Gadamer. While the former doctrine has been the major belief until the eighteenth century, the latter is recently dominating the disputes. Therefore, there are increasing attempts to learn from art and aesthetics.

Perhaps because of the advances in technology, and the fact that digital media is becoming an inevitable part of our everyday life, the idea that we can learn from aesthetics needs more attention. Udsen et al. [2] articulate that we are witnessing the “aesthetics turn” and delineates four major contributions of aesthetics in Information Technology and design of Human Computer Interaction: cultural, functionalist, experience-based, and techno-futurist. On the cultural level, we are witnessing that the computer is serving as a medium engaging us in experiencing digital media and art in our daily life. In the functionalist approach, aesthetics enables higher levels of behavioral functionalities and enhances the usability of computers. On the experience-based level, aesthetics nurtures emotions and creativity in design of interactions. The rather youngest of all other three levels, the techno-futurist approach, is the influence of the technological human reflecting on technology which in turn defines new ways of interactions between computers and humans.

Donald Norman in his book *Emotional Design: Why we love (or hate) everyday things* [5] delineates three levels of aesthetics cognition: visual, behavioral, and visceral. The visual level is the first and immediate level of aesthetics cognition; for example, when we see a beautiful flower we may appreciate its visual appeal. The behavioral level deals with the situation when we like how an artifact is functional and usable. The visceral level needs more patience; for instance, when we solve a mathematical problem, or when we contemplate an artwork and after a while we understand it (the Aha moment).

Quantifying aesthetics in visual design is more plausible than in visual arts: Unlike visual arts where the main goals may be abstract, visual design is conceptualized and created to convey a message and communicate with audiences [6–15]. This is a key but subtle difference. In fact, it is a critical measure of success in applied arts along with the visual appeal of the design. A design message has to be conveyed at first glance. Studies suggest that designers need to make a good first impression only in some few milliseconds [16, 17], and this impression deals with expressive aesthetics [18]. These qualities suggest that it is

theoretically legitimate to evaluate a visual design, and argue if it is a good design or not. In other words, an artwork may be considered as “high” or “low” art (descriptive terms from Kormsmeier [4]), but a visual design —as a very specific instance of art— may be considered as “good” or “bad”.

Although it is legitimate to distinguish a visual design as good or bad, it is much easier to identify a bad design rather than a good one. Part of the reason is the fact that there is not only one good design. For a given task in design, designers may create several alternatives. A bad design, on the other hand, is easier to distinguish, especially when it does not follow the principles of design in a professional way. On such an occasion, design critics may argue that, for instance, this design “is not visually balanced”; or “colors are not chosen carefully”. This inspires us to strive to understand what visual balance is, and to what degree a design is balanced. A similar question is raised for other elements of design, for example, color and type. In this thesis, we raise and answer some of the questions regarding quantifying aesthetics of visual design, but first, in the following section, we provide a taxonomy on quantifying aesthetics.

2.2 Taxonomy

In this section, we present a taxonomy on quantifying aesthetics based on our recollection of prior work. Figure 2.1 illustrates a schematic view of this taxonomy. To us, there are two main approaches to quantifying aesthetics: human inspection and automatic. By human inspection approach, we mean the set of empirical studies based on human judgments. This approach is well-established in psychology. The main goal is to accept or reject a hypothesis by capturing responses of participants to a set of stimuli, under very specific conditions. The automatic approach, on the other hand, tries to evaluate aesthetics through closed form mathematical expressions and artificial intelligence approaches such as pattern recognition/ machine learning and evolutionary computing.

Nevertheless, as we discuss in the following, both approaches may inspire each other. For instance, a series of studies in human inspection may lead to a ground truth for an

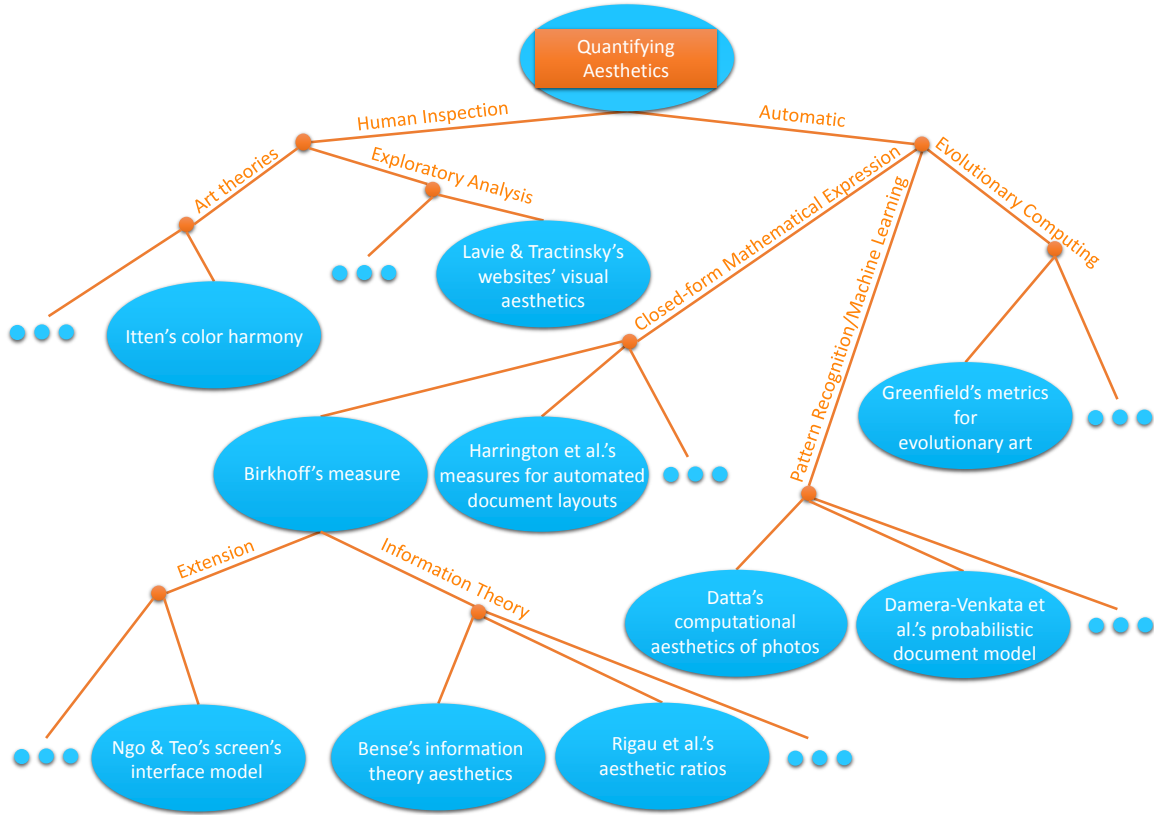


Fig. 2.1.: Taxonomy of quantifying aesthetics: A schematic view. Refer to [19–28].

automatic clustering; or a closed-form mathematical solution may steer a trend of research in empirical studies.

2.2.1 Automatic Approaches

One of the earliest attempts in quantifying aesthetics is perhaps Georg Birkhoff's measure:

$$M = \frac{O}{C}. \quad (2.1)$$

Birkhoff in his book *Aesthetics Measure* [20] argues that:

The typical aesthetic experience may be regarded as compounded of three successive phases: (1) a preliminary effort of attention, which is necessary for the act of perception,

and which increases in proportion to what we shall call the *complexity* (C) of the object; (2) the feeling of value or *aesthetics measure* (M) which rewards this effort; and finally (3) a realization that the object is characterized by a certain harmony, symmetry, or *order* (O), more or less concealed, which seems necessary to the aesthetic effect ([20], chapter 1, section 2, p.3).

Birkhoff's measure is general and implies that when complexity of a piece increases, its aesthetics value decreases. He applies this measure to quantify aesthetics of polygons, ornaments and tiles, vases, music, and poetry. Later, Bense [29] from an information theory perspective, applied redundancy theory for describing order, and Shannon's theory for complexity [30]. Also, Rigue et al. [23] define a set of aesthetics ratios based on information theory, and apply it to evaluate van Gogh's painting style [31].

Another recent extension of Birkhoff's measure belongs to Ngo et al. [21, 22] for measuring aesthetics of screen interfaces. In their model, order is a linear summation of thirteen measures such as balance, sequence, and rhythm.

From a psychological approach, Berlyin [32] studies the relation between complexity and aesthetics. Berlyin proposes a downward U-shaped graph to express the relation of complexity with aesthetics. He argues that when a piece is too simple or too complex, its aesthetics cognition decreases. A historical review on Birkhoff's measure and its branches is provided by Gary Greenfield in [30].

We enumerate some of the other closed-form mathematical expressions. Machado and Cardoso [24] present an implementation of a theory of image processing aesthetics in which image complexity is proportionally, and processing complexity, is inversely related to aesthetics. In computing aesthetics of automated documents, Harrington [33] suggests a measure as a nonlinear combination of design features in the page layout. For computing visual balance of an image, Lok and Feiner [34] suggest a measure based on lightness weightmaps computed from the image. Balinsky [35] proposes a measure of symmetry in screen interfaces. Klinger and Salingeros [36] propose a measure for evaluating aesthetics (in terms of interestingness) of a pattern based on variety and symmetry of elements in a pattern.

Automatic approaches in quantifying aesthetics are evolving by advances in pattern recognition/ machine learning and evolutionary computing in artificial intelligence. Datta et al. [28] suggest a computational model for photographic images based on a set of features such as low level color features, distinct regions in an image, spatial composition, and depth of field in photography. Damera-Venkata et al. [27] propose a probabilistic document model for automated document layouts in which some design decisions for the page layout are included. In evolutionary computing, Greenfield defines several aesthetic metrics in optimizing the achievement of evolved visual imageries [25, 37]. Ross et al. [38] also use an aesthetics model (the *Painting Bell Curve* by Ross, the first author) in synthesising evolutionary images.

2.2.2 Human Inspection Approaches

Itten [19, 39] was a color theorist who endeavored to bring color harmony from the realm of subjective to objective. Itten defines seven kinds of color contrast: contrast of hue, light-dark contrast, cold-warm contrast, complementary contrast, simultaneous contrast, contrast of saturation, and contrast of extension. Later, Mastuda extends the concept of color harmony and defines harmonious hue and tone templates [40, 41]. Kobayashi is another color theorist whose *Color Image Scale* [42, 43] in associating colors with linguistic concepts (color semantics) is well-received. There exist other well-known theories on other visual design principles. In visual balance, for instance, Arnheim quantifies a structural net that contributes to a balanced spatial composition [44]. From a psychology and neuroscience perspective, scientists have studied and validated Arnheim's net through experimental and statistical analysis (see, e.g. [45–47]).

Through exploratory factor analysis, Lavie and Tractinsky [26] propose a measurement instrument for both classic (usability and behavioral) and expressive (beauty and visually pleasing) aesthetics. Their measures are inferred from a collection of user experiments on websites. Moshagen and Thielsch [48] also delineate a set of visual aesthetic factors under four categories of simplicity, diversity, colorfulness, and craftsmanship.

2.3 Conclusion

In this chapter, we discussed the theoretical considerations on the legitimacy of quantifying aesthetics. We summarized that although some art philosophers believe that aesthetics is a matter of taste, others (and especially contemporary philosophers) argue that it is a matter of cognition and hence learning. We then presented our taxonomy of quantifying aesthetics, based on our recollection of prior work in the fields of computer science/ engineering, psychology, and neuroscience. We conclude that quantifying aesthetics is not only viable, but an urge in leveraging adoption of art and media in our contemporary life.

3. DESIGN MINING COLOR SEMANTICS

3.1 Introduction

Unlike visual arts where the main goals may be abstract, visual design is conceptualized and created to convey a message and communicate with audiences [6–15]. This is a key but subtle difference between visual arts and visual design. In fact, it is a critical measure of success in applied arts along with the visual appeal of the design. A design message has to be conveyed at first glance. Studies suggest that designers need to make a good first impression only in some few milliseconds [16, 17], and this impression deals with expressive aesthetics [18], which is a matter of visual appeal by itself.

An important design element that has a vital role in conveying a message is color. That is, in visual design, color combinations are designed for a practical purpose: not only to be appealing, but also to be silent salespersons that communicate with the audience [49]. Hence, it is important for designers to sell the right colors [50–52] to attract the observer and to inspire trust [17].

In schools of design, students spend many years to learn general principles of design, and in particular, creation and usage of color palettes to convey emotions and ideas [53]. Even within designers, there is a difference between first year and senior students [52, 53]. However, in today’s “Instagram” world, everyone wants to create designs with the purpose of self-publishing. Unfortunately, good design cannot be learned in an instant, and this presents a barrier for individuals and small to medium businesses who want to create good designs but cannot afford to hire a professional designer.

This chapter presents a first step towards our broader goal of making design accessible to the general public. We seek to understand and statistically model designers’ intuitions, and to use the knowledge gained to help non-designers in self-publishing, and designers in drawing inspiration and creativity. In particular, we focus on understanding meanings and

semantics of colors. We ask and answer the following question: by mining the work of designers, can we build a statistical model to learn color semantics? More specifically, can we build a model which associates color combinations with linguistic concepts?

Seeking good examples of visual design, we collected a relatively large corpus of magazine covers. We chose magazine covers for two reasons: First, by using a variety of data sources including the Internet, archives of our own university library, and widely accessible newsstands, we were able to create a dataset of 2,654 covers from 71 titles. Second, designers spend many days conceptualizing and creating a magazine cover that attracts the audience at first glance while competing with other magazines on a newsstand [54]. In this process, designers carefully choose a color palette for a cover based on the issue’s topic and the cover lines to convey a very specific concept. In summary, by using magazine covers we can obtain high quality examples of good design with relatively little effort and use them in our statistical model.

Although creating good design by selecting appropriate color palettes is difficult, distinguishing good and bad design is very easy even for non-designers. For instance, if a massage therapy website is designed with vivid red colors, most people will agree that it does not convey “soothing” emotions. Therefore, we validate the output of our statistical model by designing crowdsourcing experiments to answer the following question: do a majority of “casual” users agree with the associations between color combination and linguistic concepts produced by our model? In order to analyze the results of our crowdsourcing experiments, we develop a novel statistical technique inspired by click models used in web search and online advertising.

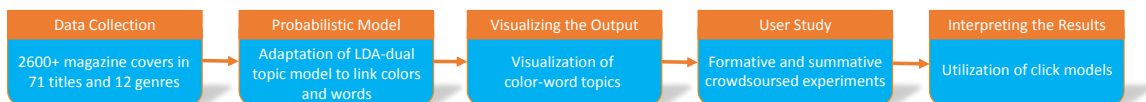


Fig. 3.1.: Overview of our end-to-end framework for design mining of color semantics.

Figure 3.1 illustrates an overview of our research process. We believe that we have developed a complete end-to-end system for semantic design mining, which includes data

collection, adaptation of a topic modeling framework, and verification of the model via a crowdsourced user study.

The flow of this chapter is as following. Section 4.2 presents a study on color semantics from both theoretical and practical aspects. This elucidates the concept of color semantics, the fine lines between color semantics and other aspects of color such as color naming and the practical advantages of color semantics in visual design. Moreover, from HCI aspect, the urge of addressing color semantics is evident, for instance, in design of metrics for quantifying aesthetics of visual designs and images, in design mining of visual designs, in design of user interactions, and in automatic design of visual media. This section is then followed by Sec. 6.3 where we review related work in computer science and engineering.

To mine designers' thought processes on color semantics, we adapt an extended version of the popular Latent Dirichlet Allocation (LDA) topic model [55], called LDA-dual [56] to jointly link the colors and words, and infer *color-word topics*. LDA-dual is originally developed by Shu et al. [56] for entity resolution in a dataset of bibliographies. Although we independently and separately developed an exact model (see Appendix B), since we later found that such a model exists under LDA-dual name, we give the credit to Shu et al. [56], and mention that we “adapt” their model. In our adaptation, the model produces combined color-word topics; the word topics are multinomial distributions on words and the color topics are histograms of colors. In Sec. 3.4 we discuss the data (designs) collection, and describe the statistical framework in Sec. 3.5.

For ease of visualization, in Sec. 3.6 we represent the word topics as word clouds, and discretize the resulting color histograms to yield a mapping to a pool of 5-color palettes. The 5-color palettes are extracted from the magazine covers by using the techniques from [57].

Unlike previous work, which has mainly focused on fitting hypotheses to crowdsourced data (i.e. human labeling) in a post-hoc manner, we conduct focused crowdsourced experiments with the sole aim of verifying whether or not users agree with the associations between color combinations and linguistic concepts produced by the model. In particular, users are shown a 5-color palette as a representative of a color topic and several choices of word clouds as representatives of word topics and asked to select the most appropriate word

cloud for the given 5-color palette. Based on the user feedback we infer the relevance of the color palette to the word cloud. We performed both a formative and a summative study and report the results. Section 3.7 describes the user studies. Utilizing a novel statistical analysis technique, inspired by click models used in web search and online advertising, in Sec. 3.8 we interpret the results of the user study. The resulting analysis shows that the results produced by our model are aligned with the intuition of the crowd.

In Sec. 3.9 we suggest a number of applications for color semantics to demonstrate how semantics can enable more meaningful user interactions, and perhaps help masses to design better. We specifically discuss color palette selection and design example recommendation, image retrieval, color region selection in images, and pattern colonization in image recoloring. We conclude this chapter and suggest future work in Sec. 6.10.

3.2 Theory

Colors are colorful: more than physical attributes of objects, colors are cognitive and have names, meanings, and values. In the following subsections, a brief review of these aspects of colors is summarized and the urge of understanding the role of color semantics in analysis and synthesis of visual designs in HCI is elucidated.

3.2.1 Color Cognition

While color perception is a rich area of research, color cognition is a relatively new area of research [58]. That is, although color is perceived as the result of wavelength discrimination, it is also cognitive [59]. Derefeldt et al. [60] elucidate that cognitive means that after visual perception color is classified to a higher level of abstraction into verbal and semantic categories [61,62] by a visual task such as color categorization.

Heer & Stone [63] suggest that when this categorization capability (in the concept of color naming) is deployed in user interfaces that model human category judgements might demonstrate more meaningful and novel user interactions. What the boundaries of color

naming are, and how this concept can be compared with color meanings has not yet been paid attention to by scientists.

3.2.2 Color Naming

The earliest effort to understanding linguistic categorization of colors is perhaps based on work of Berlin & Kay for elucidating that color naming is universal and related to evolution [64]. Before their work, most linguists believed the concept of color naming — associating names with colors— is not universal. As Kay & McDaniel [65] summarize, the two main misconceptions about color naming between linguists were that color naming is a matter of cultural relativism and semantic primes in languages are discrete entities. The latter doctrine was not able to address compound terms, e.g. *green-blue*. This inadequacy motivated Kay & McDaniel to describe color categories as continuous functions using the fuzzy set theory. The universality claim of color naming [64] pertains under some conditions. Palmer in his book *Vision Science: Photons to Phenomenology* [66] extensively explains this concept. Although color naming associates words with colors, there are other trends of research that investigate meanings and higher level of abstractions/semantics of colors.

3.2.3 Color Meanings and Semantics

The first systematic approach to quantifying “meanings” of linguistic concepts comes from *measurement of meaning* by Charles Osgood [67, 68]. Osgood proposes a semantic space based on pairs of polar terms, such as *happy-sad* or *kind-cruel*, and terms it a “semantic differential” as an “objective index of meaning” [67]. Osgood then studies cross-cultural generality of his semantic space and finds significant similarities between different primitive cultures [69]. Later, Osgood compiled a 620-concept *Atlas of Affective Meanings*, explained in [70] with a report of cross-cultural studies in 23 cultures. Adam & Osgood [71] investigate the 8 color concepts of Osgood’s Atlas —*color* (vs monochrome), *white*, *grey*, *black*, *red*, *yellow*, *green*, and *blue*— among 20 countries and report that for instance, the

relative affective meaning of *red* is *strong* and *active*. Another similar attempt belongs to Wright & Rainwater [72], yet with a more visual communication language perspective in color meanings and connotations. This is where a set of more professional color semantics such as “elegant” or “showiness” is emerging.

A breakthrough in design-based color semantics is the study by Kobayashi [42]. Kobayashi defines a meaning scale to relate “worlds of people and objects with worlds of colors” and terms it *Color Image Scale* [43]. This scale comprises of two bipolar dimensions, *warm-cool* and *soft-hard*. Kobayashi’s scale has significant contributions to visual design: it is based on 3-color combinations (3-color palettes) rather than single color patches (swatches); it also relates these 3-color combinations with two levels of abstractions, one with 180 semantics, and the other with 15 higher semantics. Because of these characteristics, Kobayashi’s scale is taught in many color courses in schools of design (e.g. see [73]). Later cross-cultural studies acknowledge the universality of Kobayashi’s bipolar scales [52, 74, 75]. Similar efforts in association of colors and words belongs to Lars Sivik as mentioned in [76].

3.2.4 Color Semantics, Emotions, and Preferences

While investigating meanings of colors, some researchers have attempted to understand the emotions and moods evoked by colors and how feelings about colors can influence our performance (e.g. see [77] and [78]). From a psychological point of view for instance, Crozier [79] argues that theories of preferences, innate and learned reactions, should be considered while studying color meanings. Crozier summarizes that *red*, for instance, has the innate mood of the alert signal and also is involved in sexual behaviour in many species; or *white* is learned to be associated with purity in some cultures. Although Crozier acknowledges that the *like-dislike* bipolar scale in prior studies is a considerable measure to research color moods, he argues that meanings of individual colors should be considered within the context (syntax, semantics, and culture) in which they are examined, also adding other factors such as age and gender.

Later, Ou et al. [74, 75, 80] studied color emotions and color preferences to clarify the relations between them. Although their studies agreed with prior work in defined scales of *warm/cool*, *heavy/light*, and *active/passive*, they found notable differences in the *like-dislike* scale between their participants, which were Chinese and British. They also found a tendency that their participants preferred color combinations which hold opposite emotions.

3.2.5 Color Semantics and Cross-cultural Considerations

Hutchings summarizes that “The Principle of Adaptation of Ideas accounts for regional variations in colour folklore. This embodies a Darwinian-type principle of behavior, that is, ‘to survive within a community a belief must have relevance to that community’” [81]. Nevertheless, it is observed that in the information era common sense about colors is increasing [82]. Such a cultural problem in HCI can be thought of as “minor science” as [83] suggests.

In a recent cross-cultural study of color emotions of 190 two-color combinations among 8 countries, *UK, Taiwan, France, Germany, Spain, Sweden, Argentina, and Iran*, Ou et al. [52] report consistency for *warm/cool*, *heavy/light*, and *active/passive*. This study also points out some inconsistency for *like/dislike*; for instance, Argentinian participants preferred more grayish colors in contrast to other participants. However, more cross-cultural studies can shed light on our knowledge in color semantics and personal preferences.

3.2.6 Color Semantics in Applied Arts

Krippendorff states that “Design is making sense (of things)” [84]. This sense making in applied arts is often equated with communication via elements of design. Professional designers and researchers have specifically emphasized on the role of colors and color semantics as a means of visual language [73, 85]. In architecture [86–88], interior design [89], textile [90], and product design in marketing [91], understanding values of colors and color semantics is acknowledged as a knowledge which can support not only non-designers but professional designers who conceptualize and ideate designs based on

their own intuitions. This knowledge can enhance the processes involved in design, such as inspiring, brainstorming, exemplifying, and communicating among designers over design prototypes (see [89]).

3.2.7 Color Semantics in HCI

Although some studies in HCI investigate the role of colors in aesthetics and usability [13, 17, 26, 48, 77, 92–103] and performance of the user [78, 104, 105], the notion of color semantics has been barely addressed. As an example, consider a scenario in which users are surfing the Web for a “massage therapy” website. Users perhaps expect to see a website with color combinations that impress them with “calm”, “soothing”, and “elegant” moods. If the color design of such a website fails to convey these moods, it will then negatively affect the trust, credibility and market of its business (See [17, 106]).

We believe that color semantics is a legitimate challenge to be addressed in various venues of HCI: in automatic design [107–112], design by example [113–116], design grammar [117], design of user interaction [118–125], quantifying aesthetics of design [26, 28, 99], user experience design [126–132], and color design [40, 133–138]. For instance, in design by example, although users are supported by a pool of designs and some mechanisms help them choose their preferred designs [116], there is no well-defined interaction to understand the user’s purpose of designs or what color combinations should be selected from their preferred designs. Accordingly, the messages that they want to communicate may be affected.

As Fallman argues, HCI needs to be understood and acknowledged as a design-oriented process in terms of philosophy and theory and methodological foundations [119]. Dearden & Finlay suggest that we need to identify patterns that are both timeless and cultural-sensitive [139]. Hoffmann & Krauss [13] conclude that many researchers do not recognize the importance of visual aesthetics in the communication intent. We need to work on semantic color vocabularies and define and propose them as domain-specific knowledge to HCI. This motivates us to revisit the concept of color semantics, utilize it in design mining

and data-driven approaches to learn from work of professional designers, and investigate theoretical and practical aspects of color semantics in user interaction design.

3.3 Related Work

Our work lies at the intersection of four research areas: color semantics and meaning, probabilistic topic models, user study via crowdsourcing, and click modeling.

3.3.1 Color Semantics and Meanings

Semantics in computer science and engineering has been observed as a gap [140–146] that merits more investigation. Nevertheless, in terms of color emotions and moods, there is a recent body of work. Csurka et al. [147] collected a dataset of color combinations and their associated labels, and applied a Gaussian mixture model on the data based on some low level color features of the color combinations. Unlike our work, their dataset is collected from a limited number of good designs suggested in [49] and mainly from an online dataset of colors [148] produced mostly by amateurs. The noisiness of the labels in this online dataset was acknowledged by both these authors and [149]. Another difference is that their model does not categorize the color combinations as joint combinations of labels and colors.

Applications of color semantics have recently come to the attention of computer engineers. In image retrieval, Solli and Lenz [150] defined a mathematical framework for Kobayashi’s Color Image Scale. This framework is utilized by [138, 151] in designing alternative and customized magazine covers for non-designers based on color moods. For color mood transfer of images, Murray et al. [152] applied the 15 moods suggested by Csurka et al. [147]. In design mining the Web, Kumar et al. [153] reported using expressive colors (color moods) for support of data-driven design tools. In contrast to our work, where we use human judgment to evaluate the goodness of our model, they use human input to build their models.

In researching linguistic categorization of colors, color naming and its cross-cultural aspects is another topic that has been under investigation for decades [66]. There is a body of work in computer science and engineering to model color naming (e.g. [154].) In recent work, Hear and Stone [63] reviewed statistical color naming models, with the goal of fitting a model to single colors and their associated names. These associations are either from human judgments or retrieved from Internet search engines. For the latter, topic modeling was utilized by Weijer et al. [155] who used Probabilistic Latent Semantic Analysis (PLSA), and Schauerte and Stiefelhagen [156] who used Latent Dirichlet Allocation (LDA). In contrast to color naming, our color semantics work takes into consideration design examples, color combinations, and different levels of abstraction. Also, unlike prior work, we adapt topic models to jointly link colors and words, and then verify the associations through user studies.

3.3.2 LDA Topic Modeling

The goal in LDA topic modeling is to infer underlying themes or topics of textual document corpora, where each topic is a multinomial distribution over words, and each document is a mixture of topics [55, 157]. Since we need to infer compound (color-word) topics for color semantics, we adapt LDA-dual [56] by viewing each magazine cover as a combined bag of words and colors. The closest work to our research is perhaps [158], where LDA is used for image annotation. The authors consider a document as a mixture of low level image features (extracted by SIFT techniques) and words. However, in our adaptation of LDA-dual model we incorporate two independent multinomial distributions, one for observed colors and one for observed words.

3.3.3 User Studies via Crowdsourcing

Viability of crowdsourcing graphical perception, including color visualization, has been confirmed by recent work of Heer and Bostock [159]. In several tests, they replicated the results of prior laboratory experiments using crowdsourcing and reported consistency

between the results. The scalability yet inexpensiveness of online experiments in colors overrides some of their limitations, such as visual acuity. For instance, Lin et al. [78] employ crowdsourcing to confirm improvement in their subjects' performance when color semantics is utilized in information visualization. Reineke et al. [99, 160] utilize cross-cultural crowdsourcing to study some perceived aesthetics aspects of visual complexity and colorfulness in a dataset of website designs. Inspired by their online setup, we designed the demographical aspects of our crowdsourcing experiments. However, we emphasize that our goals with the crowdsourcing experiments are completely different. While they wish to infer bias in perception based on demography and cultures, we validate the output of a statistical model.

3.3.4 Click Modeling

The goal of click modeling is to model users' interactions with sponsored search results or ads. In a recent work, Govindaraj et al. [161] reviewed click models, and suggested a new model by taking into account relations between the user's clicks on a list of URLs (as a response to the user's query). Modeling the user's back and forth clicks on a list of URLs, they then infer the probability of clicking on different vertical positions of URLs, regardless of their contents. Inspired by this work, we model the participants responses as a click modeling problem. Arguably our model is simpler than those proposed by existing work. However, unlike web-search where users can choose to not click on any of the results, in our case the users need to make a choice. Furthermore, since users participate willingly in the user study, the levels of user engagement are high. Therefore, we believe that our simple model suffices.

3.4 Data Collection

Our dataset of magazine covers includes 2,654 covers in 71 magazine titles and 12 genres. We collected approximately 1,500 of these covers by scanning them from magazines held by a number of libraries and newsstands in our university. The rest of the cover images

were downloaded from the Internet. Although we developed a web crawler tool to collect magazine covers, because many magazine publishers do not provide archives with high quality images, in half of the cases we had to collect online images by hand. This data will be made available for academic research upon request. We attempted to collect roughly 12 different genres of magazines to capture different contexts of design. These genres include *Art*, *Business*, *Education*, *Entertainment*, *Family*, *Fashion*, *Health*, *Nature*, *Politics*, *Science*, *Sports*, and *Technology*. To this end, we loosely followed the Dewey Classification method [162], the WorldCat indexing system [163], suggestions from our librarians, as well as the description of the magazine by the publishers. Table A in Appendix A contains a summary of our dataset. Note that the genres are tentative. In fact, this supports the use of LDA: magazines rarely include only one topic and usually are a combination of different topics (or genres in this context).

3.4.1 Preprocessing

The preprocessing of cover images was performed using a MATLAB toolkit. A color calibration process was applied to the scanned images. We use 512 basic colors which is the quantized space of sRGB with 8 bins in each channel. In this basis, the 512 basic colors are independent of each other. Given the color basis, each magazine cover (image) is then a histogram of these colors. To feed the images to LDA-dual, we scale them to 200×300 pixels using bicubic interpolation.

To extract color palettes from images we used techniques from [57], saliency map [164], and segmentation [165]. These color palettes are used for visualizing the inferred color combinations.

3.4.2 Word Vocabulary

To capture the words to be associated with color distributions of the magazine covers, the words on the covers were transcribed by hand. To create a word vocabulary, we first prune the transcribed words and then create a histogram of words. Because a more

meaningful vocabulary results in more meaningful topics, we filter out special characters, numbers, and stop words (e.g. articles and lexical words). The list of excluded words is provided in table A in Appendix A. Compound words with hyphen or dash are decomposed. A version of the Porter Stemming algorithm [166] is used to equate different forms of a word, for instance “elegant” and “elegance.” Finally, a mapping from month to season is applied. In order to include the context and classes of magazines in the associated words, the periodical category to which each magazine title belongs was added to the set of words. We collected these periodical categories from the WorldCat indexing system, which is the largest international network of library content and services [163].

3.5 Statistical Model

Ideating about visual designs, designers take into account the topic or the context in which they are asked to carry their messages. For instance, when the context is about *politics*, the designer may expose a tendency of using darker, “heavier” and “formal” colors. However this is not the only factor, the words in the design also influence a designers choice of colors. Figure 3.2 illustrates that *pink* – which is usually used for feminine related topics – has been used in a variety of magazines from different genres. This observation suggests that each design’s theme might be a combination of words and color distributions; and each design may include a proportion of various themes. Our goal is to model these combinations of words and colors, and infer proportions of these combinations in magazine cover designs. A similar intuition has been argued in statistical topic modeling, specifically LDA [55], for modeling word distributions in documents as proportions of different word topics.

LDA (Latent Dirichlet Allocation) is an intuitive approach to infer topics from text data. As Blei et al. [55, 157] describe, instead of categorizing and exploring documents using tools such as keywords, we may first categorize documents based on topics. This allows us to explore topics of interest and find related documents. For example, a document about sociology may include different topics, such as biology, evolution, history, and statistics,



Fig. 3.2.: Pink is used in all of these designs, despite the fact that each of these designs belongs to a different context and class of magazines. Images from [167–169].

with different proportions. Each of these individual topics can be viewed as a multinomial distribution over a fixed vocabulary of words. Accordingly, each document, which can be viewed as a bag of words, is a combination of these topics with some proportion. Typically, a value for the number of topics is chosen by hand. The latent topics as well as the topic proportions of each document are inferred by LDA using the observed data, which are the words in the documents.

Just like word topics are distributions over words, one can think of color topics as distributions over colors. One may argue that by only using LDA, we might be able to model the distribution of the colors on the covers, and infer the color topics. We can then utilize crowdsourcing, and ask participants to label the inferred color topics to conclude combined color-word topics, and hence the color semantics associations. The problem with this hypothesis is, we cannot capture the associations created in each design by designers. Moreover, we cannot perform this modeling in an automatic fashion since we need human judgements for labeling. Hence, we will lose the valuable insight that designers put into selecting color and word topic combinations.

A better approach is to model the links or the associations between the color topics and word topics to infer combined color-word topics. This means that the LDA model plate needs to be extended. This was recently proposed by Shu et al. [56]. The key in this

extension is the proportions vector θ . In our case, for each cover design, any proportion is a combination of color assignments as well as of word assignments. Similar to LDA, these assignments are modeled with multinomial (and conjugated by Dirichlet) distributions.

3.5.1 Review of LDA-dual Model for Color Semantics

In this section, we review the LDA-dual model proposed by Shu et al. [56] and explain how to adapt this model for color semantics. It is worth mentioning that we independently and separately derived the exact model as LDA-dual. See Appendix B for our derivations. However, since we later found that the exact model had been proposed earlier by Shu et al. [56] and had been called LDA-dual, although for a complete different problem, we decided that in this thesis we will give the credit to them and mention that we “adapt” their model. Our implementation, however, is an extension of the Matlab Topic Modeling toolbox [170] developed by [171] for LDA. Our implementation is available on request. Our implementation, however, is an extension of the Matlab Topic Modeling toolbox [170] developed by [171] for LDA. Our implementation is available on request.

Assume there are K color-word topics denoted by k_1, k_2, \dots, k_K and D magazine covers denoted by d_1, d_2, \dots, d_D . Let W denote the number of words in the vocabulary and C denote the number of color swatches, where each swatch is a patch of color defined by using its sRGB values¹. Moreover, let M_d denote the number of words and N_d denote the number of color swatches in magazine cover d_d . Let $w_{d,m}$ denote the m -th word in the d -th document and $c_{d,n}$ denote the n -th color swatch in the d -th document. Each magazine cover includes some proportion of the color-word topics. These proportions are latent, and one may use the K dimensional probability vector θ_d to denote the corresponding multinomial distribution for a document d_d . Let α , β , and γ be hyper-parameters of the Dirichlet distribution.

Given the above notation, the generative model for LDA-dual can be written as follows:

1. Draw K word topics $\psi_k \sim \text{Dirichlet}(\gamma)$.

¹Recall that we discretize and use 8 values for each of the three sRGB color channels. Therefore $C = 512$.

2. Draw K color topics $\phi_k \sim \text{Dirichlet}(\beta)$.
3. For each document $d_d \in \{d_1, d_2, \dots, d_D\}$:
 - Draw $\theta_d \sim \text{Dirichlet}(\alpha)$.
 - For each word $w_{d,m}$ with $m = 1, \dots, M_d$
 - Draw $y_{d,m} \sim \text{Discrete}(\theta_d)$
 - Draw $w_{d,m} \sim \text{Discrete}(\psi_{y_{d,m}})$
 - For each color $c_{d,n}$ with $n = 1, \dots, N_d$
 - Draw $z_{d,n} \sim \text{Discrete}(\theta_d)$
 - Draw $c_{d,n} \sim \text{Discrete}(\phi_{z_{d,n}})$

A graphical model for this generative process is illustrated in Fig. 3.3, where the shaded nodes denote observed random variables and the unshaded nodes are latent random variables.

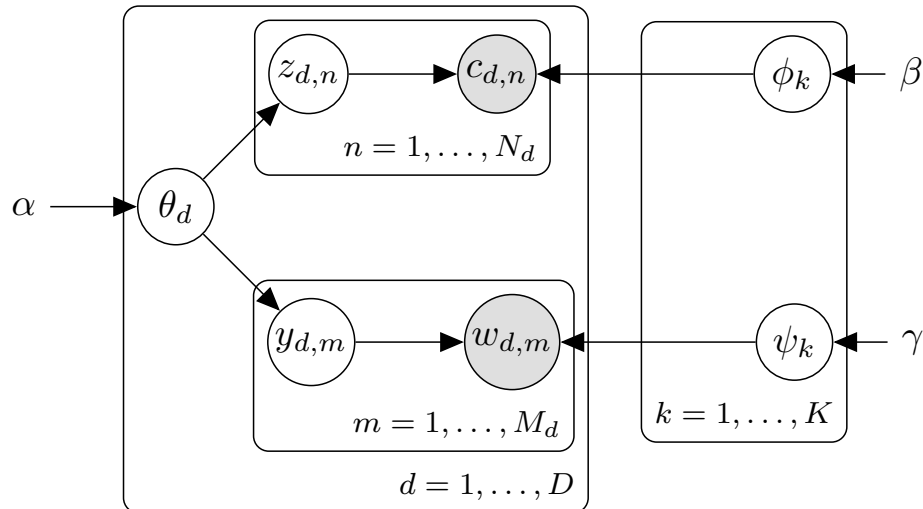


Fig. 3.3.: Hierarchical Bayesian plate model for the LDA-dual model, which combines color and word topics. D is the number of magazine covers; K is the number of color-word topics; N_d and M_d are the number of color swatches and words, respectively, in the d -th magazine cover.

If we let $\phi = \{\phi_1, \dots, \phi_K\}$, $\psi = \{\psi_1, \dots, \psi_K\}$, $\theta = \{\theta_1, \dots, \theta_D\}$, $z_d = \{z_{d,1}, \dots, z_{d,N_d}\}$, $y_d = \{y_{d,1}, \dots, y_{d,M_d}\}$, $\mathbf{z} = \{z_1, \dots, z_D\}$, $\mathbf{y} = \{y_1, \dots, y_D\}$, $\mathbf{w} = \{w_1, \dots, w_d\}$ and $\mathbf{c} = \{c_1, \dots, c_d\}$, then the joint distribution corresponding to the LDA-dual model above can be written as $p(\phi, \psi, \theta, \mathbf{z}, \mathbf{c}, \mathbf{y}, \mathbf{w}) =$

$$\prod_{i=1}^K p(\phi_i|\beta) \cdot p(\psi_i|\gamma) \cdot \prod_{d=1}^D p(\theta_d|\alpha) \cdot \left(\prod_{n=1}^N p(z_{d,n}|\theta_d) p(c_{d,n}|\phi, z_{d,n}) \right) \cdot \left(\prod_{m=1}^M p(y_{d,m}|\theta_d) p(w_{d,m}|\psi, y_{d,m}) \right). \quad (3.1)$$

Figure 3.4 provides a graphical illustration of the generative mechanism and the inference procedure described below.

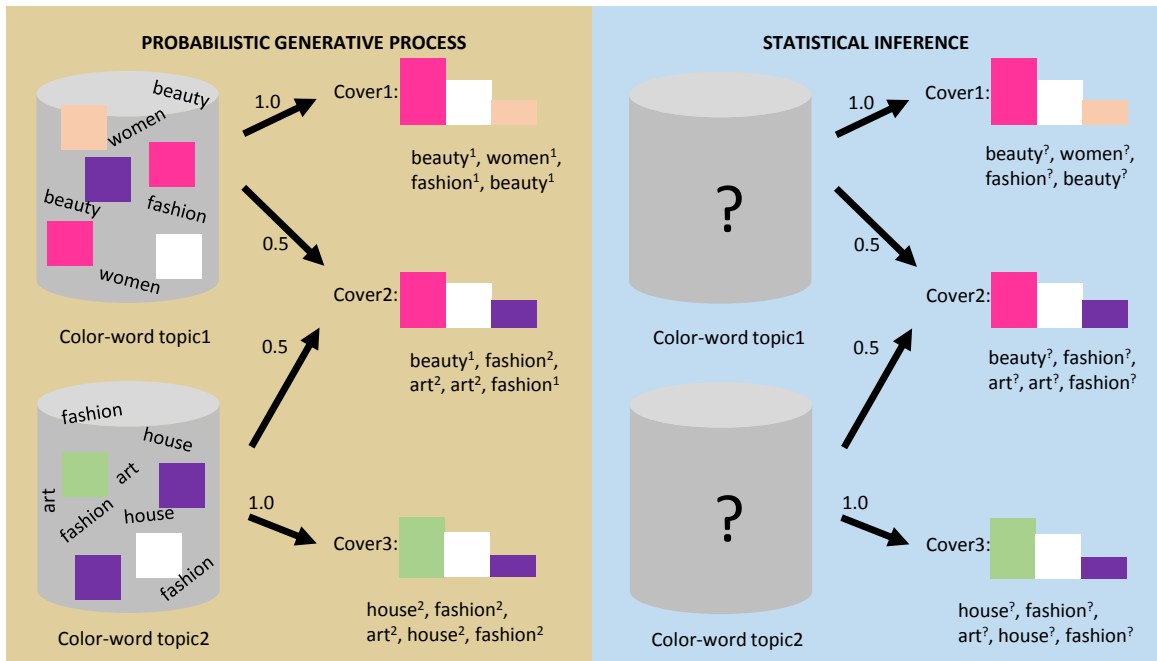


Fig. 3.4.: LDA is both a generative and inference model. This image is inspired by [172].

3.5.2 Inference

Since \mathbf{c} and \mathbf{w} are observed, inference entails computing

$$p(\phi, \psi, \theta, \mathbf{z}, \mathbf{y} | \mathbf{c}, \mathbf{w}) = \frac{p(\phi, \psi, \theta, \mathbf{z}, \mathbf{y}, \mathbf{c}, \mathbf{w})}{p(\mathbf{c}, \mathbf{w})}. \quad (3.2)$$

Theoretically, the above distribution can be obtained first by computing the joint probability distribution of the latent and the observed variables and then computing the marginal probability of the observations. In practice, however, topic modeling algorithms approximate the result to bypass the computational complexity of the solution. There are often two approaches for this approximation [157], variational inference [173, 174] and Markov chain Monte Carlo (MCMC) sampling [175, 176] (and [56, 171, 177] for collapsed Gibbs sampling). We use MCMC collapsed Gibbs sampling as implemented in the Matlab Topic Modeling toolbox [170] developed by [171] for LDA.

Although some variants of LDA can automatically find an optimal value for the number of topics K , based on our domain knowledge from the data collection process, we simply set $K = 12$. Because each color-word topic includes proportions of the color basis and the vocabulary words, we visualize a topic as a pair of colors and words histograms. Figure 3.5 illustrates the 12 topic histogram pairs for $K = 12$, $\alpha = 0.8$, and $\beta = \gamma = 0.1$. We heuristically tune these values to produce visually pleasing color histograms and semantically meaningful word histograms. The visualized histograms just illustrate the principle components. Of course, other levels of granularity can be visualized if needed. For instance, see Fig. A.4 in Appendix A for when we choose $K = 24$.

Figure 3.6 illustrates the proportions of each of the inferred color-word topics for the magazine title designs in the dataset. For instance, note that *Vogue* as a fashion magazine has k_8 and k_9 as the dominant color-word topics. As can be seen, k_8 and k_9 contain words such as “women”, “fashion”, “love”, and “beauty”, while the corresponding color histograms contain pastel and pink colors, which are often associated with fashion magazines. On the other hand, *Horticulture* which is a nature magazine has the highest proportion of k_1 ,

which pre-dominantly contains shades of green. The words in k_1 include gardening-related words such as “gardens”, “landscapes”, and “plants”.

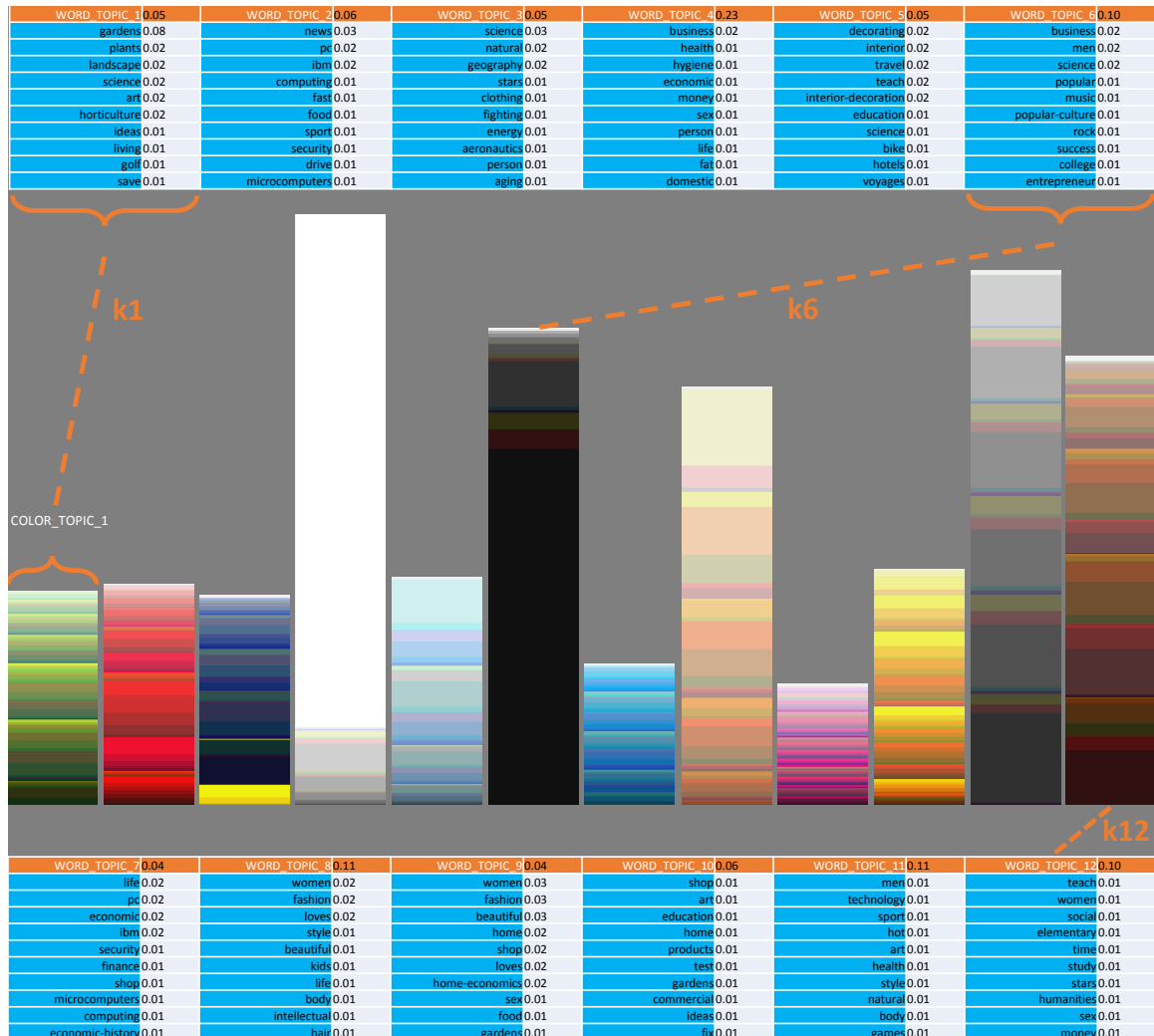


Fig. 3.5.: Color-word topics inferred by the LDA-dual model. Illustration of the 12 color topics in the middle, their corresponding 12 word topics, 6 on top for the first 6 color histograms from the left, and the other 6 at the bottom. Note that for visualization, only the principle elements in the histograms are shown. Also note that the numerical weights of the word topics are shown next to the each word topic histogram. Similarly, the numerical weights of the words in each word topic histogram are shown next to their corresponding words. See Fig. 3.6 for the proportions of the 12 color-word topics for the magazine covers in the dataset.

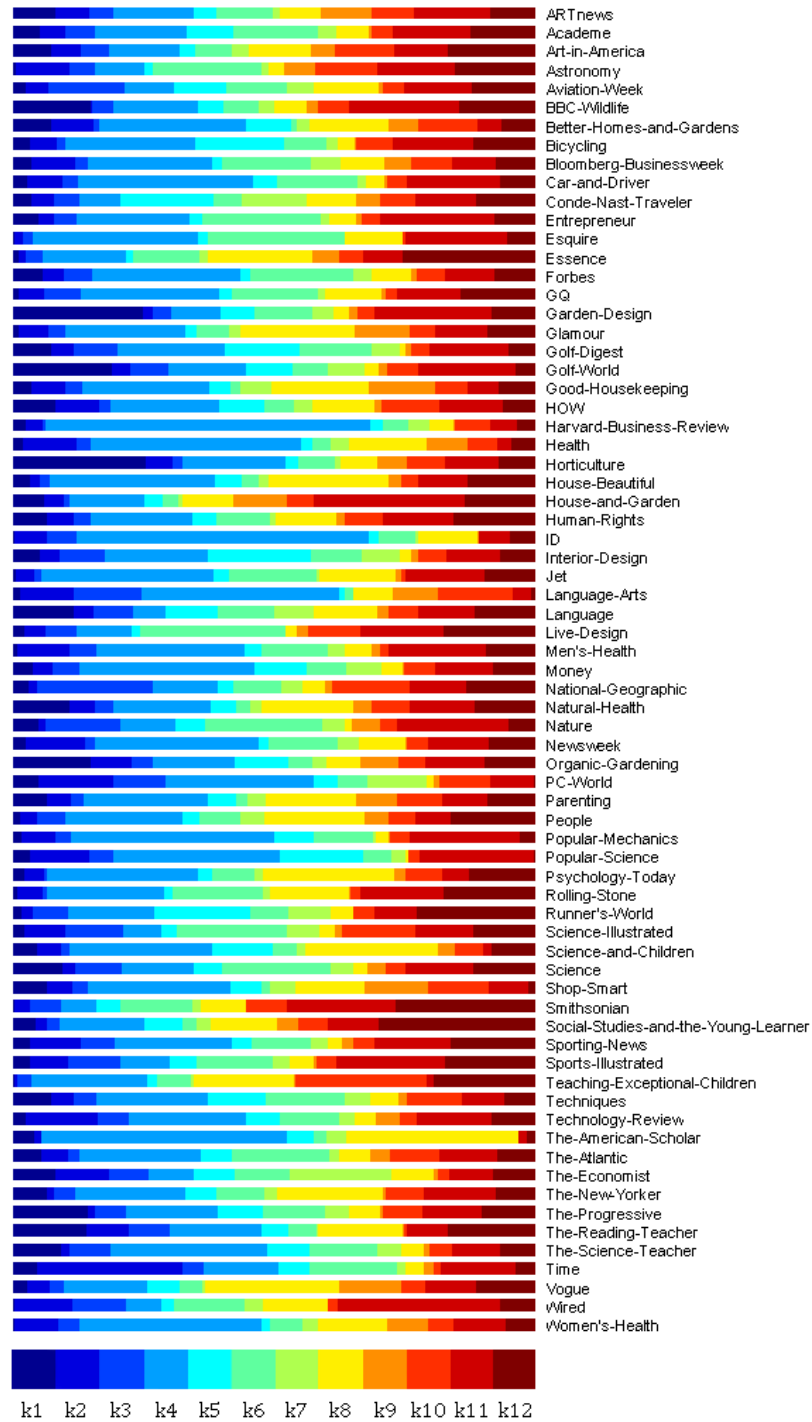


Fig. 3.6.: Topics vs. titles. The proportion of each of the 12 color-word topics, k_1 to k_{12} (see Fig. 3.5) for each magazine title is illustrated. Note that the colors here are just legends for the purpose of visualization, and not related to the color-word topics.

3.5.3 Effect of Color Basis

We chose the quantized sRGB basis of 8 bit depth for the color basis because it is simple and widely used. However, this is not the only possible choice. For instance, we can change the color basis to Kobayashi’s 1,170 3-color combinations (Color Image Scale). That is, imagine instead of 512 independent pixels in the space of colors, there are 1,170 compound 3-pixels. The difference is that the 512 basic colors in the sRGB basis are independent of each other, whereas in the case of 3-color palettes, 3 swatches (patches) of basic colors together define one color unit. For projecting images on Kobayashi’s 3-color palettes, we use the mathematical framework in ref. [138, 150]. The results using Kobayashi’s color image scale basis is provided in Fig. A.5 in Appendix A. We note in passing that different bases may lead to more insights into the data. For instance, the shades of the color topics are aligned with the color topics in the 512 color basis. Also, the words that Kobayashi has chosen are in many cases synonyms to the word topics in the 512 color basis analysis. However, we leave the deep investigation of the effect of color basis for future work.

3.6 Interpreting Model Output

Visualizing the results of LDA is a topic of research [178, 179]. Chaney & Blei [178], for example, suggest a visualization mechanism for exploring and navigating through inferred topics from LDA and their corresponding documents. Although their work does not completely address the usability evaluation of this mechanism, it inspired our visualization mechanism for our user study. In order to evaluate the color semantics hypothesis, we need to display both the color histogram and the word histogram to the participants in our user study in a comprehensive yet unbiased fashion. We address this via a two step process. The word histograms are converted to word clouds, while the color histograms are converted to color palettes using the mechanism described below. Figure 3.7 illustrates the visualization process.

3.6.1 From Color Histograms to Color Palettes

We use 5-color palettes as proxies to represent each of the color histograms in Fig. 3.5 (a). These palettes are drawn from a pool of 5-color palettes in the dataset. As mentioned before, this pool is created by using techniques from [57].

But, why 5-color palettes? The intuition behind this is that when designers are asked to design a piece such as a magazine cover, poster, or webpage, they often start by choosing a good image and then extracting the color palette from the image. The color palette is then used consistently through the design process. Designers usually choose a 3-color, 5-color, or occasionally a 7-color palette so that their designs are clean and sophisticated as opposed to busy and cluttered.

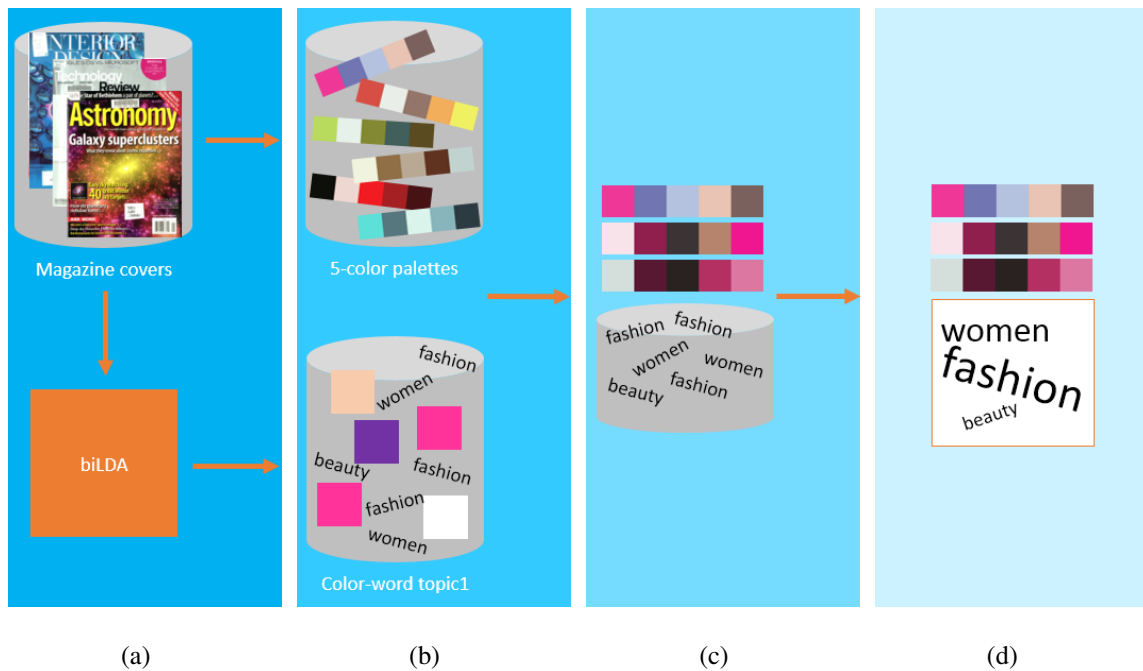


Fig. 3.7.: Visualization process for the inferred color-word topics. To visualize the color-word topics histograms inferred by the model (see Fig. 3.5), we use 5-color palettes and word clouds as proxies to color histograms and word histograms, respectively.

In order to find the closest 5-color palettes to the color topic histograms, we define a similarity metric as follows: Let S^{512} denote a color topic histogram with the 512 color

basis, and S^5 denote a 5-color palette. An intuitive similarity metric is a Euclidean distance between color swatches of S^{512} and S^5 . Among the color spaces, the CIE L*a*b* color space with a D65 reference white point is considered as a perceptually uniform space under $\Delta E \leq 2.3$, where ΔE denotes the distance between two colors, and the approximate value 2.3 is JND (Just Noticeable Difference) [180].

Defining the color similarity distance problem as a bipartite graph matching between S^{512} and S^5 with 512 and 5 nodes, respectively, we find the minimum distance cost of this graph using the Hungarian method [181]. Equation 6.3 defines the distances d_{WED} between the nodes of these two graphs. Here the weight w_i corresponds to the weight of the i -th color in the color topic histogram S^{512} . This metric can be thought of as a version of The Earth Mover’s distance suggested by Rubner et al. [182] for image retrieval, with the weight vector as the representative of color importance.

$$d_{WED} = \sum_{i=1}^{512} \frac{1}{w_i} \sum_{j=1}^5 \|S_i^{512} - S_j^5\|. \quad (3.3)$$

Computing d_{WED} for a given color topic histogram and all 5-color palettes, we choose the first 3 closest of them as proxies to the histogram (see Fig. 3.7 (c)). In the user study, we present two series of questions for the first and the second closest color palettes, because just one color palette cannot provide an adequate visualization of the whole topic histogram. See Fig. 3.9. Figure 3.8 illustrates the three first closest color palettes to a set of color histograms.

3.6.2 From Weighted Bag of Words to Word Clouds

Figures 3.7 (c) and (d) illustrate how we visualize each word topic histogram with a word cloud. Word cloud (or tag cloud) is a visualization technique used to show the relative weights of words through different font sizes. The weights resemble frequency of occurrence or importance of the words in a word dataset. A suite of word cloud algorithms and their usability is discussed in [183]. Because of the popularity of word clouds in visualizing categories and the fact that words are randomly scattered over a layout, we use

this technique in our summative user study. Using wordle², we generated black and white word clouds to avoid introducing any color bias.

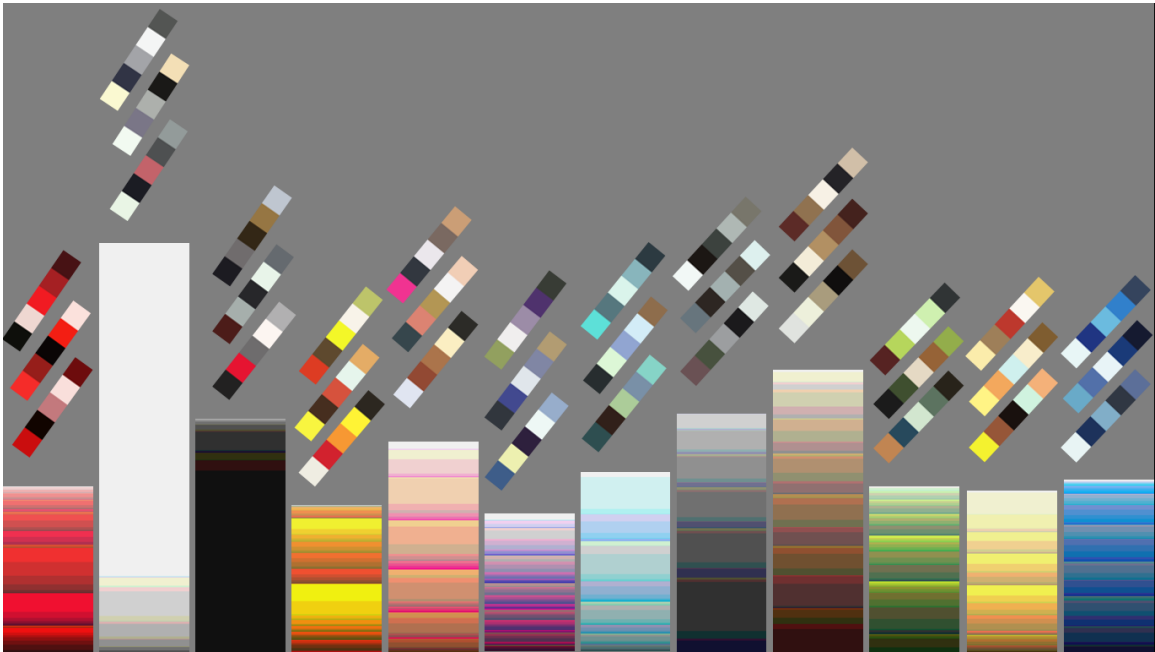


Fig. 3.8.: Three first closest color palettes to each color histogram, in the pool of color palettes. Eq 6.3 describes the used similarity metric.

3.7 User Study

The main aim of our user study is to validate the output of the probabilistic topic model. In particular, we want to understand if casual users (who are not necessarily designers) agree with the association between color combinations and linguistic concepts produced by our model. In particular, as discussed above, we represent color combinations by 5-color palettes and linguistic concepts by word clouds. We show the users a color palette and a choice of word clouds and ask them to select the most appropriate word cloud for the given color palette. Further, we describe details.

²<http://www.wordle.net>

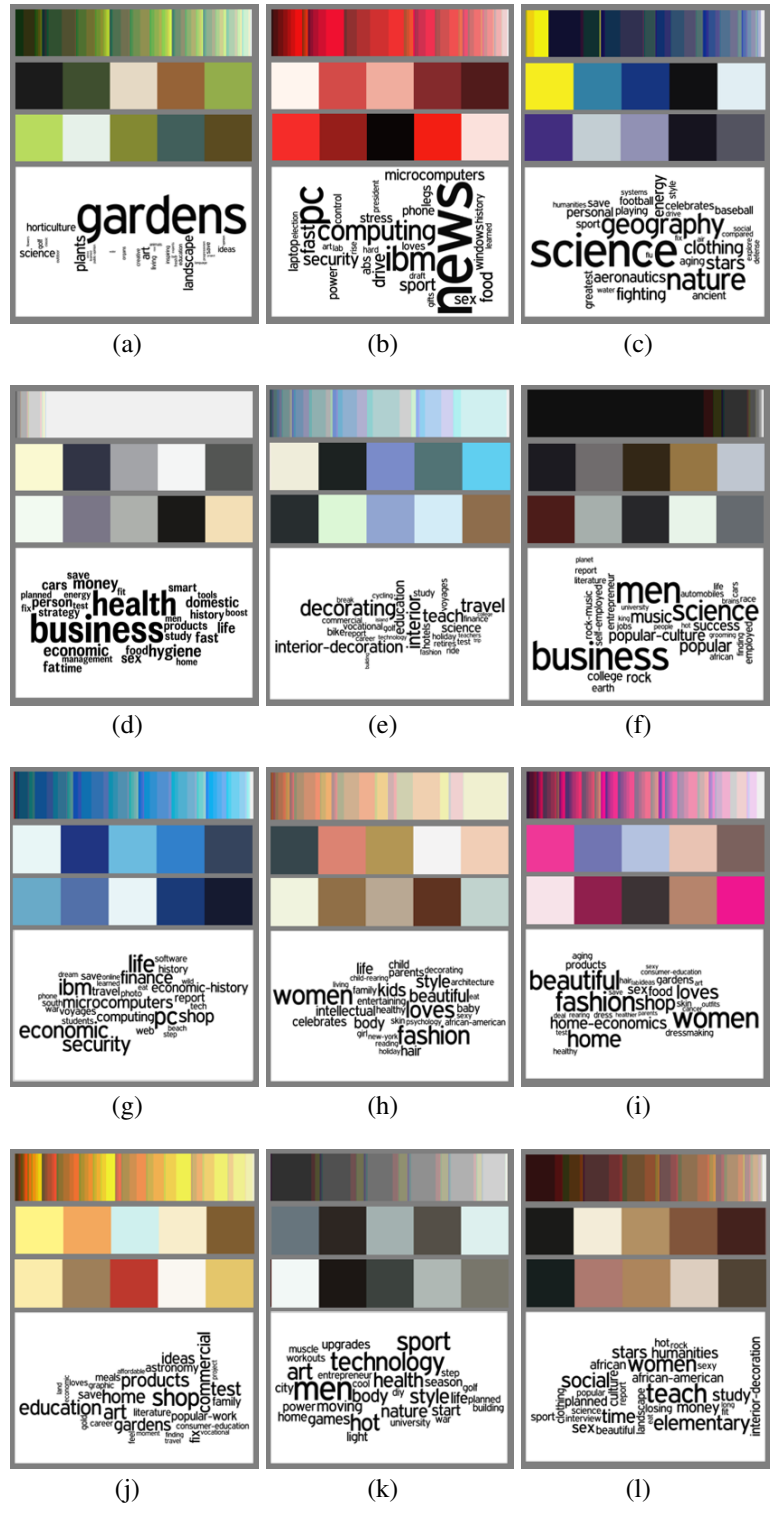


Fig. 3.9.: The 12 color-word topics produced by the model are visualized in this figure. The top panel shows the color histogram, the second and third color panels show the top two color palettes we extracted from this histogram. The word topics are visualized in the bottom panel as word clouds, with the size of a word being proportional to its weight.

3.7.1 Formative Study

The goal of this study was to learn an effective method to design questions, visualize the color-word topics, and communicate with users about color semantics. For the demography questions, we borrowed a setup similar to the one in [99]. For the color semantics, we designed two sets of 12 questions. The first and the second sets contained the 5-color palettes that were closest and second closest to the color topics, respectively. As mentioned earlier, we used 5-color palettes as proxies to color topics histograms for the sake of visualization. Each question then contained 12 multiple (check box) answers. Each answer included a triple of words – we did not use word clouds for this formative study. These triples were hand chosen by looking at each word topic histogram (see Fig. A.2 in Appendix A). Our aim was to observe the matchings that users perform between the color palettes and the word triples.

This study was designed as an online survey³ hosted on our university’s survey platform. The survey was advertised through social networks and our university’s network. A summary of what we learned from 367 responses follows: In most of the questions, participants showed extremely high agreements to the colors and words associations inferred by the model. However, we conjectured that this could be because the three representative words were chosen by us, whereas a more unbiased method would be to just let the participants select the concepts from the word topics histograms. This was further supported by comments from a number of participants who argued that they were not able to comprehend some of the word triples. Some participants also mentioned that iterating through a list of 12 word triples for each question was rather tedious. Another feedback was about the demography questions, some participants stating that they did not like such detailed questions in a free survey.

³<http://goo.gl/WFnjHL>

3.7.2 Summative Study

Participants

Our color semantics online survey⁴ is hosted at our university's survey system. It was advertised through social networks and the university email network. So far, we have collected 859 responses from 487 (56.69%) females, 367 (42.72%) males, and 5 others (0.58%), in the age range of 18 to 80 years (with mean = 30.98). The participants are from 70 countries with the majority from the US (59.84%) and are native speakers of 66 languages, . There are 348 (40.51%) participants who have lived in more than one country. There are 352 (40.97%) participants with college degrees, 451 (52.50%) with graduate degrees, and 55 others (pre-high school, high school, and professional degrees). The majority, i.e. 716 (83.35%) participants are non-designers. In contrast, there are 130 (15.13%) participants with three or more years of experience in visual design (including graphic design, interior design, and textiles.) Additional statistics are reported in Fig. A.3 in Appendix A.

Stimuli and Procedure

Figure 3.10 illustrates the flow of the survey. In order to simulate a matching experiment between pairs of color and word topics, we designed a question as follows: one 5-color palette is shown on the left side of the screen, and three shuffled and random word clouds, as well as a “None of the above” option (multiple choice check boxes) are shown on the right side in a vertical order. Among the three randomly drawn word clouds, one is the word cloud inferred from the model. The reason for dividing the survey into two subsets of questions is because otherwise participants lose interest in finishing the experiment. This observation was made in the early stages of the experiment. The reason for shuffling the word clouds in each question is that we conjectured there might be a position bias in the vertically arranged options. In Sec. 3.8, we explain how we were inspired by click models in [161] and utilized this framework in analyzing the experimental results.

⁴<http://goo.gl/P4W9XL>

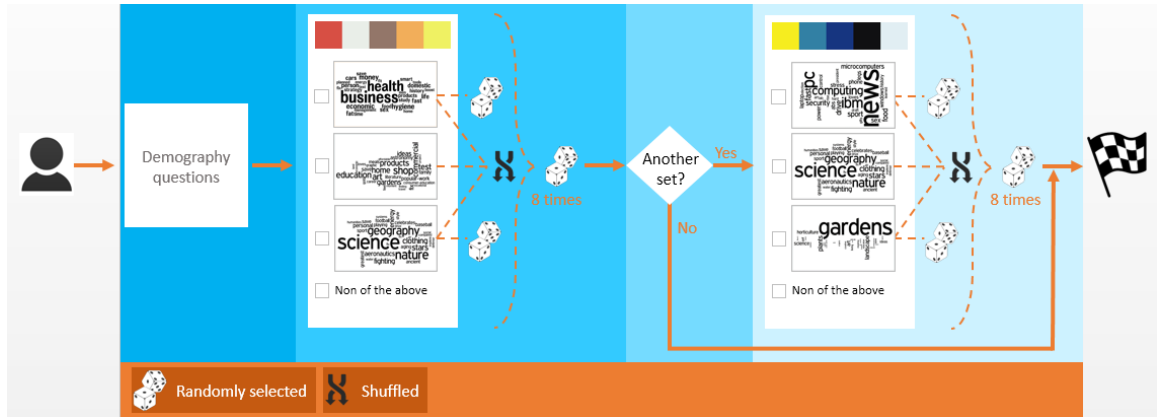


Fig. 3.10.: Flow of the summative user study.

We created 24 questions for the first and second closest 5-color palettes corresponding to the 12 inferred color topics. However, to avoid exhausting the participants, we draw 8 questions randomly and ask the participants to answer them. Then we ask the participants if they would like to continue by taking another set of 8 questions. Of all participants, 61.35% of the users chose to continue and answered all 16 questions.

3.8 Interpreting the User Study

In this section we explain the statistical inference mechanism that we used to understand the user responses.

3.8.1 Statistical Model

First we define some notation. Let c_i denote the event that the i -th color palette was displayed. Also, let w_j denote the event that the user selected (clicked on) the j -th word cloud, and u_{ij} denote the probability that j -th word cloud was selected (clicked on) by the user in response to the i -th color palette. In order to compute u_{ij} we note that

$$u_{ij} = \Pr(w_j | c_i). \quad (3.4)$$

There are three possible positions $p \in \{1, 2, 3\}$ at which a word cloud can be displayed. Let d_{jp} denote the event that the j -th word cloud was displayed at position p , and let w_{jp} denote the event that the user selected the j -th word cloud which was displayed at the p -th position. Then

$$u_{ij} = \sum_{p \in \{1, 2, 3\}} \Pr(w_{jp} | d_{jp}, c_i) \cdot \Pr(d_{jp} | c_i). \quad (3.5)$$

If d_j denotes the event that the j -th word cloud was selected for display and d_j^p the event that it was displayed at position p , then

$$\Pr(d_{jp} | c_i) = \Pr(d_j | c_i) \cdot \Pr(d_j^p | c_i). \quad (3.6)$$

According to our experiment design, each word cloud has an equal probability of appearing in any of the three positions. Therefore

$$\Pr(d_j^p | c_i) = \frac{1}{3}. \quad (3.7)$$

On the other hand, we always select the i -th word cloud (the true word cloud as per our model) for the i -th color palette. The other two slots are filled by selecting any two of the remaining 11 word clouds uniformly at random. Therefore

$$\Pr(d_j | c_i) = \begin{cases} 1 & \text{if } i = j \\ \frac{2}{11} & \text{otherwise.} \end{cases} \quad (3.8)$$

All that remains is to estimate $\Pr(w_{jp} | d_{jp}, c_i)$. For this task, we borrow from the cascade click model [161] and write the probability as a product of the following two factors:

- the probability that the p -th position is examined by a user, the so-called position bias. It is denoted b_p .
- the intrinsic relevance of the word cloud j to the color palette i . This is the quantity that we seek to infer from the user responses, and we will denote it r_{ij} .

In other words,

$$\Pr(w_{jp} | d_{jp}, c_i) = r_{ij} \cdot b_p, \quad (3.9)$$

and by using (3.5) and letting $\sum_{p \in \{1,2,3\}} \Pr(d_{jp}|c_i) \cdot b_p$ be q_{ij} we can write

$$u_{ij} = r_{ij} \cdot \sum_{p \in \{1,2,3\}} \Pr(d_{jp}|c_i) \cdot b_p = r_{ij} \cdot q_{ij}. \quad (3.10)$$

Note that b_p can be pre-computed as follows:

$$b_p = \frac{m_p}{m}, \quad (3.11)$$

where m denotes the total number of trials (each question in our survey is equivalent to one trial), and m_p denotes the number of times the word cloud at position p was selected in any of the trials.

Let m_i denote the number of trials in which the i -th color palette was displayed, and m_{ij} denote the number of trials in which the i -th color palette was displayed and the j -th word cloud was selected. We can assume that the trials are independent, and therefore the probability of observing this data under model (3.10) can be written as

$$\Pr(m_i, m_{ij}) = (r_{ij} \cdot q_{ij})^{m_{ij}} (1 - r_{ij} \cdot q_{ij})^{m_i - m_{ij}}. \quad (3.12)$$

The maximum likelihood estimate for $r_{ij} \cdot q_{ij}$ is simply $\frac{m_{ij}}{m_i}$, from which we can infer \hat{r}_{ij} , the maximum likelihood estimate for r_{ij} , as

$$\hat{r}_{ij} = \frac{m_{ij}}{m_i \cdot q_{ij}}. \quad (3.13)$$

\hat{R}^1	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9	w_{10}	w_{11}	w_{12}
c_1	2.14	0.33	1.39	0.39	0.49	0.38	0.26	0.38	0.31	0.81	0.43	0.30
c_2	0.57	0.28	0.56	0.62	1.18	0.38	0.35	1.76	1.87	0.71	0.68	1.45
c_3	0.14	1.27	1.53	0.84	0.89	0.89	1.18	0.54	0.30	0.67	1.04	0.34
c_4	0.03	1.62	1.13	1.28	0.71	1.19	1.68	0.12	0.28	0.63	1.43	0.30
c_5	0.38	1.23	1.44	1.09	1.45	1.05	1.61	0.43	0.36	0.93	1.12	0.53
c_6	0.74	0.99	1.08	0.74	0.53	1.85	0.99	0.26	0.10	0.59	1.02	0.40
c_7	0.13	1.33	1.23	1.03	1.12	1.01	1.63	0.29	0.48	0.38	1.04	0.48
c_8	0.51	0.33	0.61	0.41	1.49	0.40	0.34	1.79	1.32	1.01	0.46	1.07
c_9	0.43	0.03	0.10	0.30	1.30	0.07	0.10	1.96	2.07	0.61	0.34	1.18
c_{10}	1.05	0.25	0.73	0.57	1.80	0.34	0.18	1.53	0.99	1.43	0.61	1.17
c_{11}	0.30	1.68	0.93	1.41	0.48	1.23	1.52	0.23	0.23	0.74	1.38	0.37
c_{12}	1.01	0.49	1.56	0.93	0.89	1.53	0.63	0.57	0.63	1.08	1.06	0.59

(a)

\hat{R}^2	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9	w_{10}	w_{11}	w_{12}
c_1	2.23	0.24	1.45	0.48	1.07	0.48	0.56	0.51	0.29	1.02	0.49	0.45
c_2	0.04	0.97	0.77	0.80	0.75	0.46	0.41	1.09	1.54	0.58	0.90	1.07
c_3	0.31	1.17	1.18	1.04	1.38	0.96	1.29	0.60	0.67	0.85	1.39	0.48
c_4	0.19	1.19	0.91	1.31	0.97	1.17	1.52	0.28	0.43	0.80	0.92	0.63
c_5	0.65	0.93	1.56	0.93	1.31	0.84	0.79	0.48	0.56	1.11	1.40	0.74
c_6	0.19	1.33	1.16	0.79	0.46	1.80	1.29	0.18	0.24	0.53	1.38	0.37
c_7	0.05	1.19	0.97	1.22	1.01	1.39	1.38	0.25	0.71	0.71	1.20	0.36
c_8	1.22	0.85	1.80	0.91	1.39	0.92	0.76	0.61	0.64	1.26	0.83	0.88
c_9	0.26	0.09	0.13	0.31	0.92	0.09	0.13	1.65	2.42	0.47	0.18	1.21
c_{10}	0.74	0.40	0.60	0.70	1.98	0.35	0.44	0.95	0.82	1.31	0.58	1.01
c_{11}	0.09	1.63	0.99	1.11	0.57	1.20	1.41	0.05	0.00	0.63	1.44	0.10
c_{12}	0.52	0.42	1.21	0.96	0.86	0.66	0.62	0.90	0.99	0.91	0.49	1.10

(b)

Fig. 3.11.: Relevance matrices \hat{R}^1 and \hat{R}^2 for the first and second set of questions, respectively. The elements of these matrices are estimated intrinsic relevance of associations between colors and words, calculated from the participants' responses. The higher the value, the greater the intrinsic relevance associated by the users. Ideally, the diagonals should contain the highest values.

3.8.2 Analyzing the Results

Figure 3.11 illustrates two relevance matrices \hat{R}^1 and \hat{R}^2 . \hat{R}^1 (resp. \hat{R}^2) corresponds to the inferred relevance of the first (resp. second) closest color palette to the word cloud produced by LDA-dual. The rows are color palettes, as proxies to color topics histograms, and the columns are word-clouds, as proxies to word topics histograms. The (i, j) -th elements of these matrices are the intrinsic relevance values \hat{r}_{ij} , computed from the observed responses of the participants using the model described in the previous section. Higher values of \hat{r}_{ij} mean that the users found a high correlation between the i -th word cloud and the j -th color palette. If the participants find the word cloud produced by our model to be the most relevant for a given color palette, then the diagonal entries, marked in blue, should contain the highest values. Whenever an off-diagonal entry is larger than the corresponding diagonal entry, it is marked yellow in the figure.

As can be seen from the relevance matrices in Fig. 3.11, in most cases the diagonal elements are higher than the off-diagonal elements. This indicates a strong correlation between the results of LDA-dual and the opinion of the participants. We studied the entries with the highest relevance values such as \hat{r}_{11}^1 and \hat{r}_{99}^1 . Referring to Fig. 3.9 (a) and (i), one reason these values are high could be because green for “garden” (\hat{r}_{11}^1) and pink and purple for “fashion” and “women” (\hat{r}_{99}^2) are intuitive and widely accepted correlations across different cultures.

There are a few color palettes such as c_2 in \hat{R}^1 where the users assign higher relevance to other word clouds than the one produced by the LDA-dual model. To understand this, note that in Fig. 3.9 (b) the first 5-color palette, which is c_2 , predominantly contains shades of red and black. Unsurprisingly, users assign higher relevance to word clouds w_8 and w_9 which are about “sex” and “beauty.” In our dataset however, the red and black color combinations are often used by news magazines such as *Time* and *The Economists*, and computer magazines such as *PC Magazine*. One can perform a similar analysis for other color palettes such as c_{12} in \hat{R}^1 and c_8 in \hat{R}^2 to infer why there is a mismatch between the model output and the relevance values assigned by the users. This also shows why it

is important to perform a user study; domain specific color palettes and their associated linguistic concepts may not always transfer to a general context.

To understand the differences between female and male participants as well as between non-US and US participants we computed the corresponding relevance matrices. See Fig. A.1 in Appendix A. Comparing these matrices with Fig. 3.11 we do not observe any significant differences. This indicates that our results do not depend significantly on the demography or cultural background of the users. However, when we compare the designers to the non-designers we note that there are more zero values in the off-diagonals for designers. This indicates that designers exhibit a stronger bias against selecting certain word clouds for certain color palettes, perhaps because of their training.

3.8.3 Association Directionality

Our interviews [184] with designers suggest that it is often easier for people to associate words with color combinations than color combinations with words. To examine the directionality of the associations, we set up another crowdsourcing experiment similar to the summative experiment. This time, however, we showed a word cloud and asked the user to match it with the shown color combinations⁵.

So far, we have collected 255 responses, with 38% male, 62% female, and 3% others, in the age range of 18 to 66 years (with average 24.34). 79% of the participants are from the US. 87% of the participants are non-designers. Figure 3.12 illustrates the relevance matrices of the responses. Generally, the values of these matrices are lower than of the matrices in Fig. 3.11. This indicates that participants were more uncertain in associating color combinations to a given word cloud. Nevertheless, still the diagonal entities in most of the cases have the highest values, suggesting that participants agreed with the associations inferred by the model.

⁵<http://goo.gl/O3Xlp2>

\hat{R}^1	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}	c_{11}	c_{12}
w_1	2.09	0.50	0.18	0.26	0.00	0.37	0.08	0.36	0.17	0.40	0.08	0.35
w_2	0.17	0.60	1.43	1.72	1.72	1.05	1.22	0.92	0.49	0.16	1.80	0.72
w_3	1.86	0.30	1.51	0.62	1.05	0.88	0.92	0.58	0.00	0.56	0.75	0.84
w_4	1.03	0.57	1.18	1.36	1.15	0.92	0.82	0.89	0.32	0.80	0.86	0.58
w_5	1.22	1.02	0.61	1.03	1.27	0.34	0.99	1.38	1.34	1.76	0.37	0.48
w_6	1.30	0.41	1.04	1.30	1.28	1.62	1.04	0.43	0.09	0.61	1.30	0.72
w_7	0.79	0.17	0.98	1.36	1.30	0.88	1.58	0.33	0.25	0.45	1.67	0.75
w_8	0.16	1.62	0.42	0.57	0.69	0.09	0.29	1.83	1.74	1.62	0.17	0.09
w_9	0.00	1.39	0.90	0.70	0.27	0.20	0.55	0.82	1.78	0.68	0.36	0.56
w_{10}	1.00	0.42	0.90	0.38	0.71	0.37	0.49	0.56	0.94	1.71	0.30	0.47
w_{11}	1.14	0.56	1.27	0.59	0.87	0.70	1.24	0.41	0.09	0.70	1.61	1.06
w_{12}	0.75	1.30	0.96	0.65	1.42	0.33	0.50	1.63	1.71	1.37	0.42	0.77

(a)

\hat{R}^2	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}	c_{11}	c_{12}
w_1	2.04	0.13	0.42	0.00	0.27	0.29	0.28	0.40	0.13	0.56	0.28	0.28
w_2	0.34	1.02	1.15	1.32	0.70	0.80	1.71	0.56	0.12	0.84	1.08	0.76
w_3	2.02	0.45	0.97	0.95	0.94	0.69	1.36	1.16	0.50	1.11	0.90	0.92
w_4	1.52	0.85	0.70	1.13	1.47	1.57	1.12	0.75	0.26	1.05	1.03	0.86
w_5	1.69	0.54	0.43	0.85	1.20	0.40	1.05	0.81	1.13	1.75	0.54	0.99
w_6	0.93	0.36	1.37	0.61	1.20	1.82	1.00	0.83	0.12	0.53	1.24	0.22
w_7	0.71	0.56	1.26	0.94	1.14	1.12	1.24	0.98	0.28	0.27	1.27	0.28
w_8	0.70	1.41	0.88	0.45	0.78	0.00	0.82	0.65	2.03	1.21	0.00	1.49
w_9	1.03	1.58	0.92	0.54	1.07	0.51	0.12	0.12	2.34	0.96	0.23	0.55
w_{10}	2.12	1.07	0.56	0.42	0.80	0.42	0.42	1.26	0.70	1.41	0.29	0.86
w_{11}	1.03	0.97	0.83	0.75	1.09	1.02	1.00	0.50	0.50	1.52	1.33	0.69
w_{12}	0.47	0.57	1.35	0.24	0.96	0.85	0.57	0.36	1.48	1.81	0.39	1.09

(b)

Fig. 3.12.: Results of examining the words to colors association directionality. Relevance matrices \hat{R}^1 and \hat{R}^2 for the first and second set of questions, respectively. The elements of these matrices are estimated intrinsic relevance of associations between words (rows) and colors (columns), calculated from the participants' responses. Compare with matrices in Fig. 3.11.

Another note to take into account is that some participants mentioned that in some cases their personal preferences interfered with their decisions in the matching processes. That is, for a given word cloud, they sometimes preferred to choose one color combination over another. This suggests that while color design is a professional task, it also needs to account for customization and preferences.

3.9 Applications

Similar to the applications of color naming proposed by Heer and Stone [63], in order to illustrate how color semantics enables more meaningful user interactions, we present a number of applications in color palette selection and design example retrieval, image retrieval, and recoloring images using semantics.

3.9.1 Color Palettes Selection Using Semantics

One way to make design accessible to the masses is to recommend design elements and design examples. In content creation, for instance, the user needs to use a color combination that is both appealing and aligned with his/her purpose of design. While this need seems more immediate to a non-designer, a designer may also prefer to see examples for a more creative or even an inspiring color combination. Our inferred color semantics can be applied in color palette recommendation. Although there are pools of color palettes available online [148, 185], it is not easy to navigate, find, or recommend a set of palettes based on the user's need (see [149] and [186]). Consider a scenario in which the user is interested to find a color palette that conveys meanings of both "technology" and "fashion." Figure 3.13 illustrates such a scenario, where the user queries for a combination of these two concepts. This type of query is mapped to the word topics, and from them to their corresponding color topics. We then map these color topics to a set of color palettes in a pool of color palettes, created from the magazine dataset using techniques in [57]. The user can then choose his/her preferred color palettes from the recommended set.

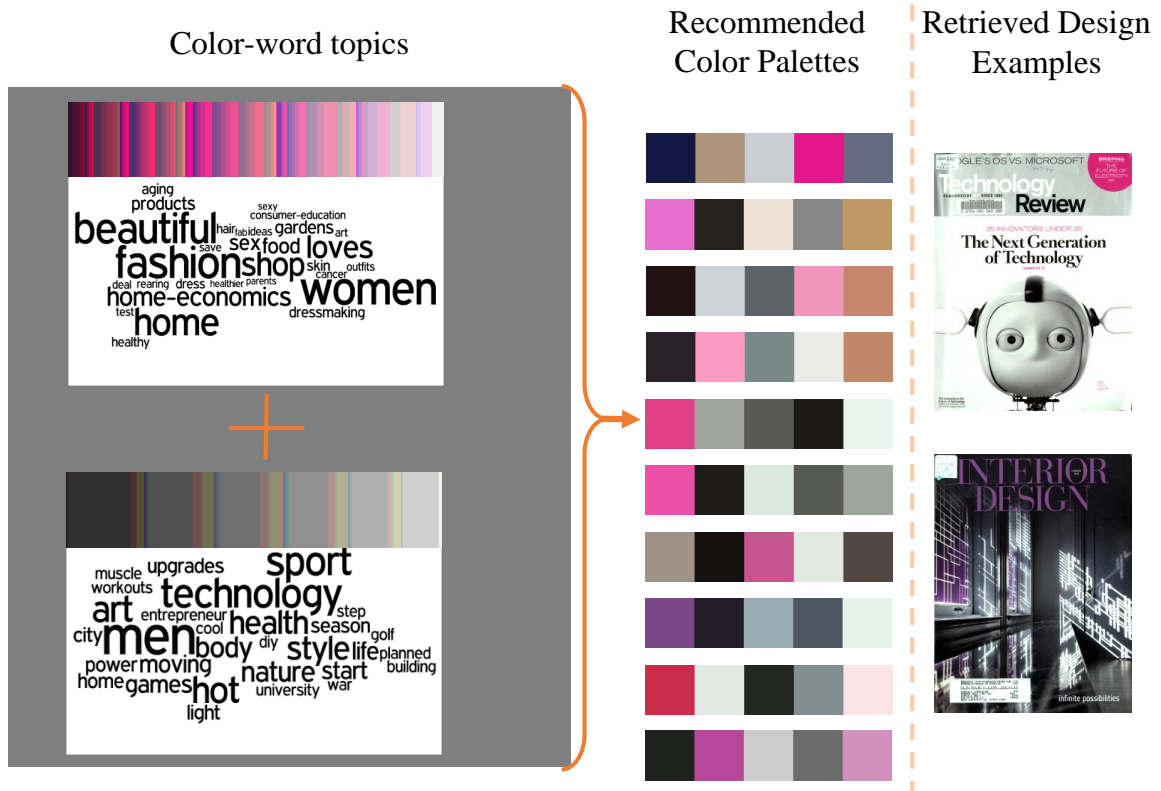


Fig. 3.13.: Application of color semantics in color palettes selection, and design example retrieval. See Applications section. Magazine cover images from [168, 187].

3.9.2 Design Example Recommendation

Another plausible application in using our color semantics is recommendation of alternative designs. This can be done in several ways, i.e. by using a given a set of recommended color palettes (see Fig. 3.13), or directly, just by taking color semantics as the user's input. Providing such examples can be utilized in creativity support tools [188]. It can also facilitate design prototyping [189]. Similarly, translating the associations inferred from the magazine covers to other kinds of media such as website remains for future work. Because LDA-dual is a generative model, it can be used in automatic creation of designs based on the user's preferences, similar to magazine cover design [151], and website design [116]. Utilizing generative models can also leverage the design-driven studies in HCI [117].

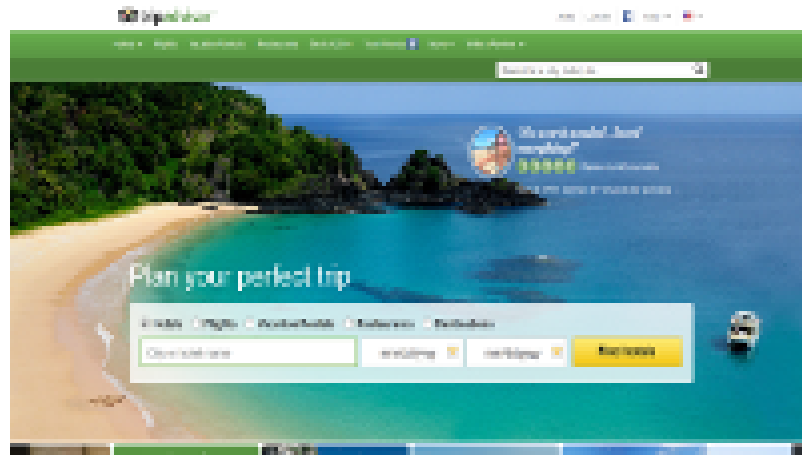
3.9.3 Image Retrieval Using Color Semantics

The gap of color semantics in computer vision, specifically for image retrieval, has been emphasized by prior work [140, 142, 143, 145]. Nevertheless, the associations of color combinations and linguistic concepts can be utilized in designing high level image features for both current image retrieval algorithms [150] as well as deep learning algorithms [190, 191]. Consider a scenario in which the user first makes a query about “interior design” in an image search community. As a result, a number of images will be retrieved and suggested. However, in order to explore and navigate through the retrieved images, the user may query for “classy” images. In this case, we are able to map this query to the color-word topics, and rank the already retrieved images based on the color topic histograms that represent “garden”.

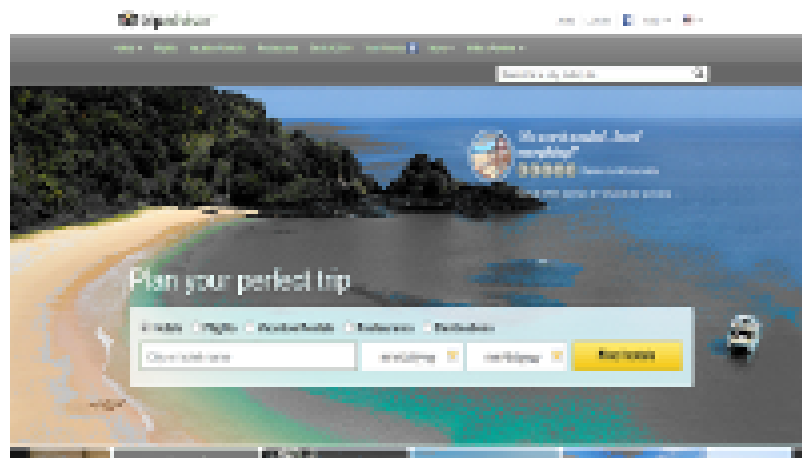
3.9.4 Image Color Selection Using Semantics

As Heer and Stone [159] note, a common interaction in image editing is to find a region of colors. They suggest two types of color region mechanisms based on color names: color name queries from the user, and using the *magic wand* tool. In our case, we use the user’s query and present the set of pixels that semantically contributes to the colors of an image. Figure 3.14 illustrates this kind of interaction. Figure 3.14 a) is the original image, a screenshot of a traveling agency website [192] nominated as one of the best designs (based on people’s votes) according to the Webby Awards 2014 [193]. We are interested to know what color regions have contributed to “travel” and “shop” in this image. Figure 3.14 b) represents these regions, while turning the other regions to the grayscale mode. In contrast, Fig. 3.14 c) illustrates only color regions that contribute to “women” in our color semantics space.

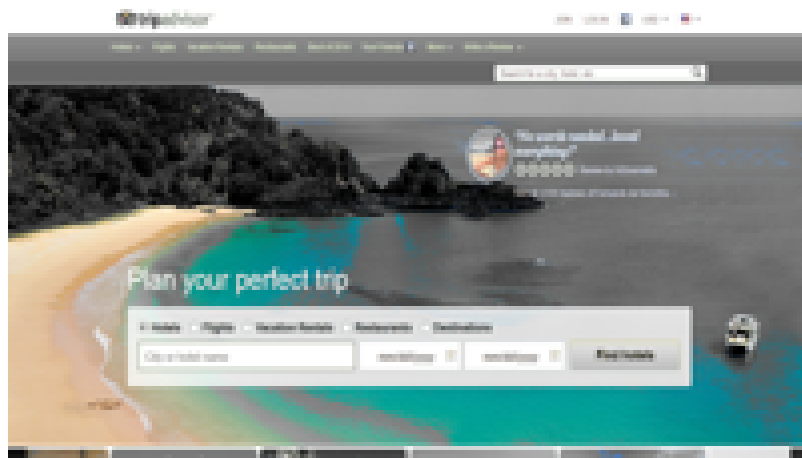
Note that in order to find these color regions we map the user’s queries to the word topics, and preserve their associated color topics in the image. Of course, utilizing the current image segmentation techniques leads to a smoother pixel selection.



(a)



(b)



(c)

Fig. 3.14.: Image color selection using color semantics. a) The original image, b) colors that contribute to “travel” and “shop”, and c) “women” and “fashion”, in the original image. Image from the homepage of tripadvisor.com [192].

3.9.5 Image Recoloring Using Semantics

Given the fact that colors communicate messages, designers often modify color themes of their designs. Part of this color modification may involve recoloring an image [194], transferring a color theme to an image [152], or enhancing a color theme of an image [195]. In this section we suggest an application of color semantics for colorizing patterns (Fig. 3.15), based on techniques introduced by Lin et al. [194].

Figure 3.15 a) illustrates a grayscale pattern. Using the color semantics associations, we are able to accept a user query, map it to the word topics, and use a 5-color palette as a representative of this color topic to colorize the original pattern. Figure 3.15 b), c), and d) illustrate the results of colorizing the original pattern using “science”, “shop”, and “sports” queries, respectively.

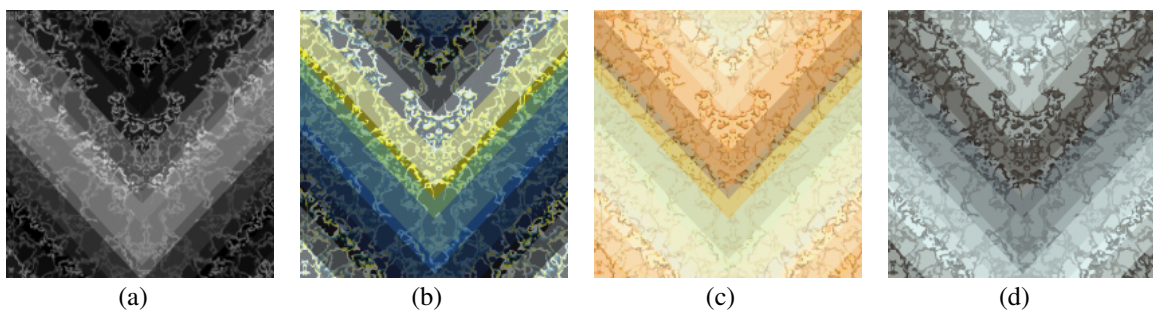


Fig. 3.15.: Pattern colorization using color semantics. a) The original pattern, b) “science”, c) “shop”, and d) “sports” and “men” colorized patterns. Images colorized using techniques in [194]. Pattern from *ColourLovers ArrayOfLilly* [148].

3.10 Conclusion and Future Work

Taking into account that the goal of visual design is both to convey a message and to appeal to audiences aesthetically, we investigate design mining of designers’ thought process in associating colors with linguistic concepts. We collected high quality examples of good designs, a dataset of 2,600+ magazine cover designs in 71 titles and 12 genres. We

then adapted LDA-dual, an extension of the popular LDA topic model, to simultaneously explain designers' both choice of colors and words on the magazine covers. To verify the inferred color semantics with the wisdom of the crowd, we used a crowdsourcing experiment. The results confirm that our model is able to successfully discover the association between colors and linguistic concepts. This completes the loop of our design mining system, from hypothesis inference to validation.

Our user study shows that users across different countries, gender, and age groups largely agree with the colors and linguistic concepts associations produced by our model. However, to put this in a context note, our experiment is biased towards an audience which is capable of reading and understanding English, is mainly college educated, and has access to the Internet. Arguably, the Internet is globalizing design by melting cultural perceptions. Talking about the effect of globalization in art, Carroll [82] argues that:

What we are witnessing now differs from the past insofar as what we see emerging is something like a single, integrated, cosmopolitan institution of art, organized transnationally in such a way that the participants, from wherever they hail, share converging or overlapping traditions and practices at the same time that they exhibit and distribute their art in internationally coordinated venues. And this, I submit, is something worth considering as substantially unprecedented.

Nevertheless, studying color semantics in different cultures and folklore (similar to [160] in studying aesthetics of low level features of colors) can increase our understanding of customized designs and hopefully help both designers and non-designers in communicating with their audience.

We then suggest a number of applications for color semantics to demonstrate how semantics can enable more meaningful user interactions, and perhaps help masses to design better. We specifically discuss color palette selection and design example recommendation, image retrieval, color region selection in images, and pattern colonization in image recoloring. Nevertheless, at this stage, these applications are not final tools. They need to be investigated more in terms of user experience design, and to be evaluated as to what degree

they can actually facilitate user interactions. Incorporating these applications in current design creation tools is also part of future work.

The present work is a first step towards our broader goal of making design accessible to the general public. We believe that this work can open up a number of interesting avenues for future work, some of which are currently being discussed in the two next sections.

3.10.1 Visual Design Language and User Interaction

From a broader perspective, association of linguistic concepts with design elements, specifically colors, can accelerate the development of a visual design language. This can enhance the user interaction mechanisms to be more intuitive and cognitive. Also, it can promote the user's involvement in design tasks such as prototyping, collaboration, and critique. The LDA-dual model can be utilized to associate the critiques (in a design critique system such as [196]) and semantic tags (in a design query system such as [153]) with design elements of the corresponding designs. This may provide more insight into understanding design critiques and queries.

3.10.2 Quantifying Aesthetics

Current computational models for quantifying aesthetics are based on low level features of design elements [28, 99, 197, 198]. We believe that both outcomes of this work, the inferred color semantics and the application of the LDA-dual model to color semantics can leverage the current models of quantifying aesthetics. While the first outcome is intended to provide more high level features to evaluate a design based on the message that it is meant to convey, the second outcome can be used in linking linguistic concepts with colors, spatial composition (layout), and typography of visual media design.

4. DESIGN MINING VISUAL BALANCE

4.1 Introduction

Psychological studies show that visual balance is an innate concept for humans [199, 200], which influences how we perceive visual aesthetics and cognize harmony [201]. There exists a body of work endeavouring to understand visual balance and its relation with symmetry [202] about vertical [203–205] and horizontal [205] axes, content of the scene [206], color contrast [207,208], and styles in abstract and representational artworks [45, 46, 209, 210].

In visual design, for instance, balance is also a key principle that helps designers to convey their messages [211–219]. Photographers, specifically, create visual balance in the spatial composition of photos through photo cropping [47, 220]. Our motivation in this work is to model visual balance. Learning visual balance from the work of professionals in design and photography may help to enable the automatic design applications in layout creation [109, 110, 112, 151, 221–234], content retargeting [235–239], cropping [220, 238, 240, 241], photo composition [238, 242] and quantifying aesthetics of layouts [22, 28, 48, 198, 243–245]. Nevertheless, there is no rigorous model to describe visual balance. In the current studies, our reference to this notion is mainly based on art theorists’ speculations and general guidelines from professional designers. However, because of access to large-scale datasets, we might be able to revisit such a theoretical concept in art and attempt for a more quantifiable definition.

In prior work, visual balance is defined as “looking visually right” [246] and is studied under the “theory of rightness in composition” [45, 47]. Balance is considered in two general categories: symmetry and asymmetry [34], which in any case relates to harmony [247]. One of the central theories around balance is perhaps Arnheim’s structural net [44], in which he hypothesizes that there are nine hotspots (including the center) on any visual

artwork, and identifies their locations. Although in prior work, Arnheim’s net is studied through psychophysical experiments, by asking participants opinions about special arrangements of visual elements in paintings and photos, to the best of our knowledge, the present work is the first design mining framework on large scale datasets of images for evaluating Arnheim’s theory.

In this chapter, we examine Arnheim’s hypothesis about balance through design mining a dataset of 120K images. These images are obtained from a professional photography website, 500px [248], and have at least 100 user likes (some of these images have several thousands of likes). Because visual balance is stimulated by the interaction of visual elements such as lines, color, texture, and orientation in an image, we run our design mining on the images’ saliency maps. This decision is justified according to the fact that this same set of visual elements comprises the underlying features in the saliency map models. Having computed the saliency maps of the images, we then model these maps as a mixture of Gaussians by utilizing GMM and EM techniques. In our modeling, we initially position the Gaussians in the same places that Arnheim locates his visual balance hotspots on an image. We describe the adaption of GMM and the scalability considerations in processing large scale datasets in our framework.

Our inferred Gaussians align with Arnheim’s hotspots, and confirm his structural net, specifically the center and the symmetry of the net. Because the dataset contains semantic tags (e.g. *landscape*, *architecture*, *fashion*, etc.) of the images, for future work, we aim to cluster some general visual balance templates, and to establish their linkage to the tags. This may lead to recommendations for visual layout of photos and visual design in general.

The flow of this chapter is as follows. In Sec. 4.2, we first discuss the theories behind visual balance, and how we intuitively model it. We then describe our formal definition and modeling framework in Sec. 4.3. The results of our work are presented in Sec. 4.4. We conclude this chapter in Sec. 4.5 with a general discussion and future work.

4.2 Theory

4.2.1 Visual Balance in Spatial Composition

Visual balance is often studied along with the spatial composition of an image. Spatial composition in an image is defined as the arrangement of visual elements and the way that they interact with each other in the space. Perhaps one of the earliest attempts in quantifying spatial structure of visually appealing artworks belongs to Adolf Zeising through defining *Golden Section* ratios [249]. Gustav Fechner later studied the Golden Section ratios, and argued that this notion is overemphasized [250]. Some studies argue against [251] the Golden Section concept, and some to the favor [252] of it. This could be because of the way that the experiments were set up [253]. Another well-known rule in creating visually appealing spatial composition is called the *Rule of Third*. While some argue to the favor of this rule, others discuss about its triviality [254]. The notion of visual balance is another component of the spatial composition has remained a challenge. In studies of balance to this date, this concept has almost always been equated with the physical connotation of balance and equilibrium (e.g., see [47]). However, visual balance to art experts might metaphorically mean harmony [247]. This inspires us to study this notion beyond a measure of weight along the vertical or horizontal axes. We aim to revisit Arnheim’s structural net, and examine it through mining from a large-scale dataset of highly liked images. We attempt to answer the following question: In this large-scale dataset, can we infer any pattern to support Arnheim’s speculation about his hypothetical hotspots in his proposed structural net? In short, our modeling framework indeed supports Arnheim’s net.

4.2.2 Visual Rightness

In prior work, visual balance is defined as “looking visually right” [246] and is studied under the “theory of rightness in composition” [45,47]. One of the central theories around balance is perhaps Arnheim’s structural net [44]. Figure 4.1 illustrates Arnheim’s net, in which he hypothesizes that there are nine hotspots (including the center) on any visual

artwork, and identifies their locations. When visual weights are located on these hotspots, visual stability and balance are more perceived.

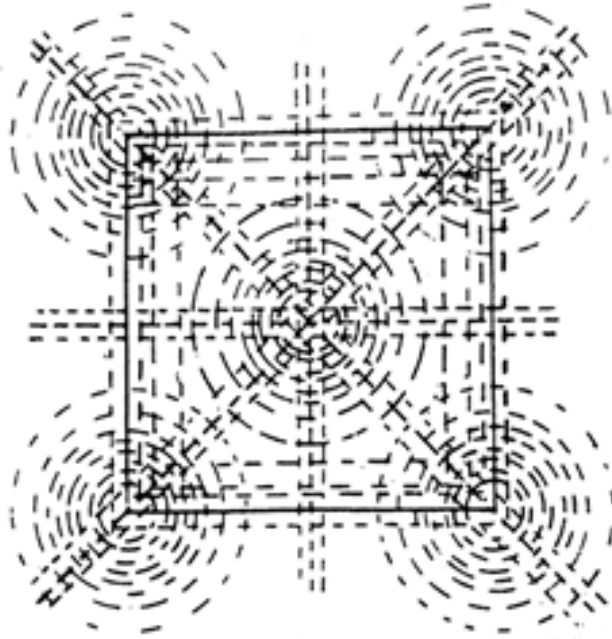


Fig. 4.1.: Arnheim's structural net. This net illustrates the hotspots on any visual artwork. Image reproduced from [44].

According to Arnheim, factors that influence visual weight are: dynamic, position, depth, size, color, intrinsic interest, isolation, shape simplicity, shape orientation, and knowledge of the scene [44]. Some of these factors are studied in prior work, e.g. features that influence dynamic quality of static abstract designs [255].

In our analysis, we model the distribution of the visual weight by elliptical shapes, more precisely by Gaussian distributions. To find the visual weights, we use the notion of saliency. This decision is justified according to the fact that this same set of visual elements comprises the underlying features in the saliency map models. Figure 4.2 illustrates an image (The Starry Night by Vincent van Gogh), its saliency map (using the algorithm described in [164]), and the overlap of salient hotspots with Arnheim's net. In Sec. 4.3, we explain how our model benefits from saliency maps.

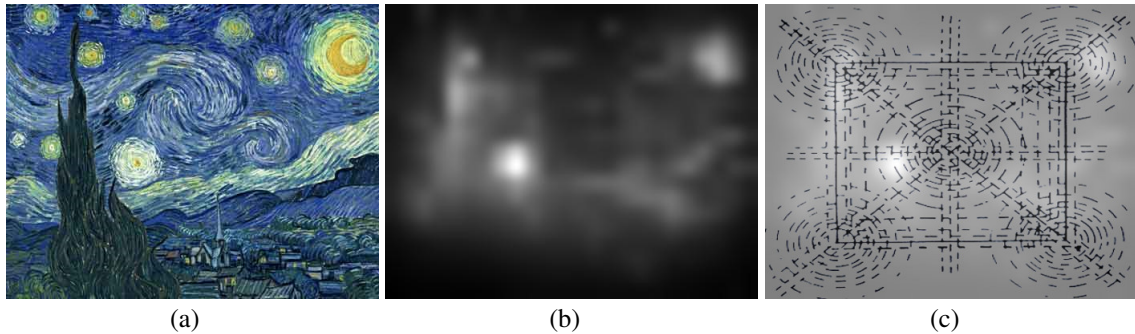


Fig. 4.2.: An example for overlap of saliency map with Arnheim's net.

4.3 Modeling Framework

A saliency map (e.g. Fig. 4.2.b) of an image is a grayscale image with pixels corresponding to the pixels in the original image. Each saliency pixel can be represented by a vector $(x, y, v)^\top$, where x and y correspond to the spatial location of the pixel, and $v \in 0, 1, \dots, 255$, the luminance value of the pixel, corresponds to the saliency value of this pixel. In other words, a higher value of v represents a more salient pixel.

Our goal is to model the saliency values of a saliency map by a mixture of Gaussians. We therefore represent a saliency map as a scatterplot of points in the 2D Cartesian space, and fit a Gaussian mixture model to the density distributions of this scatterplot. We define a saliency scatterplot of a saliency map as follows. For each pixel $(x, y, v)^\top$ in the saliency map, we generate v numbers of point $\mathbf{x} = (x, y)^\top$ in its corresponding saliency scatterplot. We denote a saliency scatterplot with N number of points in set $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$. In this fashion, we can represent the value of a saliency pixel as a measure of density of points in the corresponding saliency scatterplot. For the Gaussian mixture analysis, we follow the notations in [256]. The Gaussian mixture distribution for point \mathbf{x} in its saliency scatterplot can be written as a linear combination of Gaussians in the form

$$p(\mathbf{x}) = \sum_{k=1}^K \pi_k \mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k), \quad (4.1)$$

where the mixing coefficient π_k is modeled as the probability of assigning pixel \mathbf{x} to Gaussian component k . To model the mixing coefficients, we introduce a K dimensional binary

random variable \mathbf{z} drawn from a multinomial distribution, where $p(z_k = 1) = \pi_k$. Hence, this distribution can be written in the form

$$p(\mathbf{z}) = \prod_{k=1}^K \pi_k^{z_k}. \quad (4.2)$$

Because π_k is a probability, it must satisfy

$$0 \leq \pi_k \leq 1 \quad (4.3)$$

together with

$$\sum_{k=1}^K \pi_k = 1. \quad (4.4)$$

Intuitively, the advantage of defining the mixing coefficients as probabilities is that we can introduce a K dimensional binary random variable $\mathbf{z} = (z_1, z_2, \dots, z_K)^\top$, where dimension z_k is the label or assignment of point \mathbf{x} to Gaussian component k with some probability π_k . This variable, \mathbf{z} , is latent to us, and by defining a joint distribution of it and observed point \mathbf{x} , we can infer the assignments using the Bayes formula. Formally, we define the joint distribution of \mathbf{x} and \mathbf{z} as

$$p(\mathbf{x}, \mathbf{z}) = p(\mathbf{x} | \mathbf{z}) \cdot p(\mathbf{z}). \quad (4.5)$$

Note that we can compute $p(\mathbf{x})$ by marginalizing this joint probability over \mathbf{z} . Because \mathbf{z} is latent, we can infer it based on the observation of point \mathbf{x} using the Bayes formula

$$\gamma(z_k) \equiv p(z_k = 1 | \mathbf{x}) = \frac{p(z_k = 1) p(\mathbf{x} | z_k = 1)}{\sum_{j=1}^K p(z_j = 1) p(\mathbf{x} | z_j = 1)} = \frac{\pi_k \mathcal{N}(\mathbf{x} | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\sum_{j=1}^K \pi_j \mathcal{N}(\mathbf{x} | \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)}. \quad (4.6)$$

In fact, in our modeling, π_k are priors for $z_k = 1$, and $\gamma(z_k)$ is the corresponding posterior after observation of \mathbf{x} .

So far, we have defined the problem for one point. Our problem for our dataset of saliency scatterplots is as follows. Denote \mathcal{X} as our dataset of I number of saliency scatterplots. We assume there are K Gaussian components underlying in each scatterplot, with mean $\boldsymbol{\mu}_k$ and covariance $\boldsymbol{\Sigma}_k$; however, each component has a different mixing coefficient

π_{ik} for the i -th saliency scatterplot and the Gaussian mixture component k . To fit a Gaussian mixture to a saliency scatterplot in dataset \mathcal{X} , we denote the i -th saliency scatterplot by $\mathbf{X}^{(i)}$, for $i = 1, 2, \dots, I$, and represent it as an $N \times D$ matrix, where N denotes the number of the points in the i -th saliency scatterplot, and $D = 2$ denotes the dimensions of each point. Note that N for each saliency scatterplot is different; however, we do not index it for simplifying the notations. In this notation, the n -th row in matrix \mathbf{X} is point \mathbf{x}_n given by \mathbf{X}_n^\top . If we assume that the points in a saliency scatterplot are independent, then the log of the likelihood function for all the I saliency scatterplots stored in \mathcal{X} is given by

$$\ln p(\mathcal{X} | \boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \sum_{i=1}^I \sum_{n=1}^N \ln \left\{ \sum_{k=1}^K \pi_{ik} \mathcal{N}(\mathbf{x}_n^{(i)} | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \right\}. \quad (4.7)$$

By maximizing the log likelihood function in (4.7) with respect to the means $\boldsymbol{\mu}_k$, we obtain

$$\boldsymbol{\mu}_k = \frac{1}{N_k} \sum_{i=1}^I \sum_{n=1}^N \gamma^{(i)}(z_{nk}) \mathbf{x}_n^{(i)} \quad (4.8)$$

where

$$N_k = \sum_{i=1}^I \sum_{n=1}^N \gamma^{(i)}(z_{nk})$$

and

$$\gamma^{(i)}(z_{nk}) = \frac{\pi_{ik} \mathcal{N}(\mathbf{x}_n^{(i)} | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\sum_{j=1}^K \pi_{ij} \mathcal{N}(\mathbf{x}_n^{(i)} | \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)}. \quad (4.9)$$

By maximizing the log likelihood function with respect to the covariance matrixes $\boldsymbol{\Sigma}_k$, we obtain

$$\boldsymbol{\Sigma}_k = \frac{1}{N_k} \sum_{i=1}^I \sum_{n=1}^N \gamma^{(i)}(z_{nk}) (\mathbf{x}_n^{(i)} - \boldsymbol{\mu}_k) (\mathbf{x}_n^{(i)} - \boldsymbol{\mu}_k)^\top. \quad (4.10)$$

Finally, we obtain the mixing coefficients π_{ik} for saliency scatterplot $\mathbf{X}^{(i)}$, by maximizing the log likelihood function with respect to π_{ik} while taking account of the constraint (4.4) using a Lagrange multiplier:

$$\pi_{ik} = \frac{N_k^{(i)}}{N} \quad (4.11)$$

where

$$N_k^{(i)} = \sum_{n=1}^N \gamma^{(i)}(z_{nk}). \quad (4.12)$$

4.3.1 EM for Gaussian Mixtures

Given the Gaussian mixture model for our dataset, we optimize to maximize the likelihood function in (4.7) with respect to its parameters $\boldsymbol{\mu}_k$, $\boldsymbol{\Sigma}_k$, and π_{ik} . The EM algorithm is then:

1. Initialize $\boldsymbol{\mu}_k$, $\boldsymbol{\Sigma}_k$, and π_{ik} , and compute the initial value of the log likelihood.
2. **E step.** Compute $\gamma^{(i)}(z_{nk})$.
3. **M step.** Recompute the estimation of the parameters using the current $\gamma^{(i)}(z_{nk})$.
4. Compute the log likelihood again, and check for the convergence condition. If the condition is satisfied, stop; otherwise return to step 2.

4.3.2 Dataset

Our dataset includes about 120K images. These images are obtained from a professional photography website, 500px [248], and have at least 100 user likes (some of these images have several thousands of likes). These images also have semantic tags (e.g. *landscape*, *architecture*, *fashion*, etc.) which enable us, for future work, to cluster some general visual balance templates and to establish their linkage to the tags. The preprocessing of the images is performed with Matlab Imaging and Parallel Computing toolboxes. This includes image resizing and computing of the saliency maps.

4.4 Results

We performed our algorithm on the Purdue Clusters using Matlab Parallel Computing toolbox. We assigned 96 cores in 24 nodes for the parallel computing. For our experiment,

we consumed about 120 hours for the EM algorithm. We initiated two types of Gaussians for our experiment, five Gaussians as illustrated in Fig. 4.3.a and nine Gaussians in Fig. 4.3.c. We initiated these Gaussians to text Arnheim's hotspots (compare with Fig. 4.1.

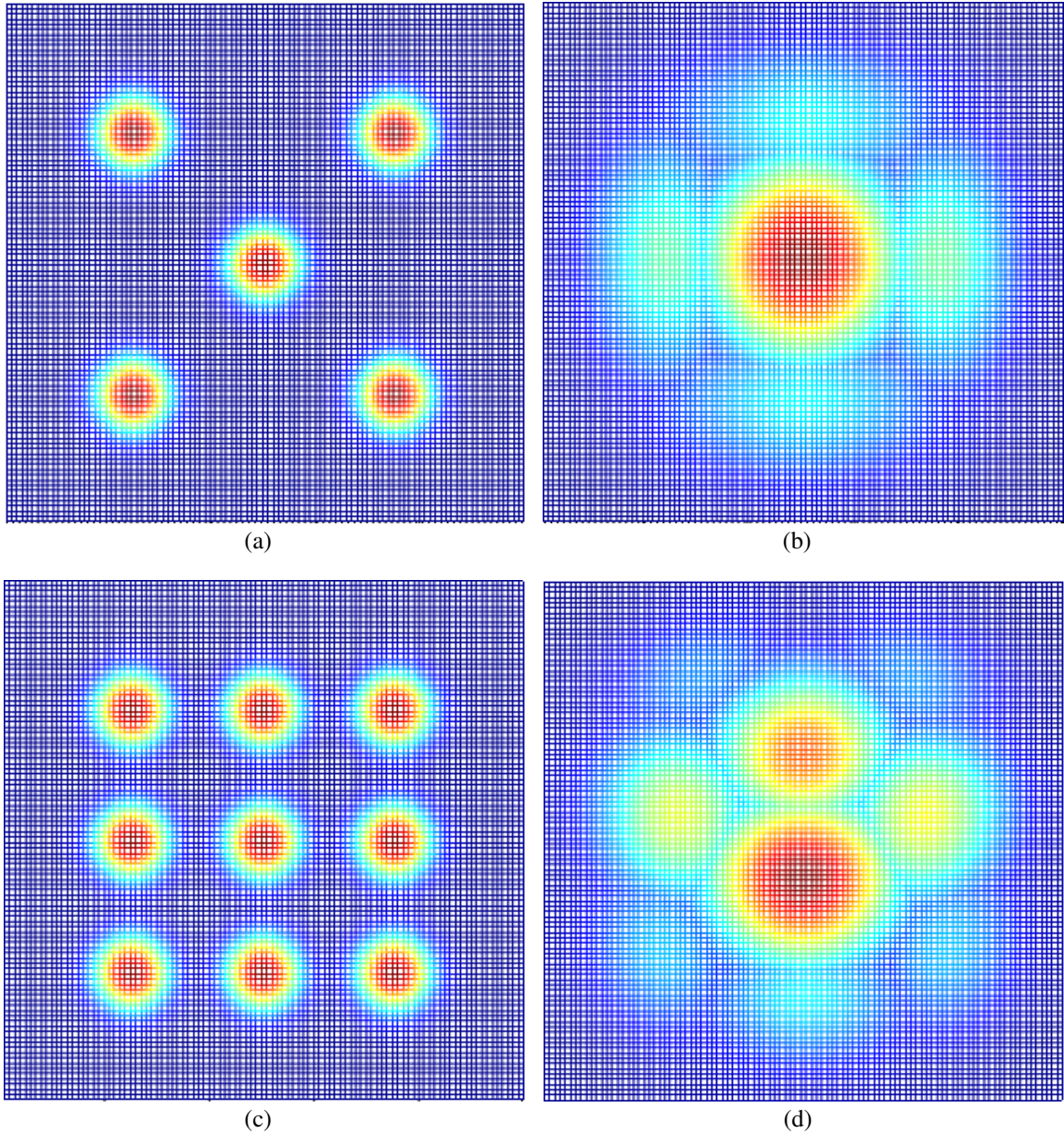


Fig. 4.3.: Results of our GMM fitting.

The results of our computations are illustrated in Fig. 4.3.b and Fig. 4.3.d for initial Gaussians in Fig. 4.3.a and Fig. 4.3.c, respectively. Our inferred Gaussians align with

Arnheim's hotspots, and confirm his structural net, specifically the center and the symmetry of the net.

Moreover, the second column of Fig. 4.3 suggests that Arnheim's power of the center [257], or the center of the mass is the most important location in the images. This supports the idea of computing the center of the mass as a means of quantifying balance. It is also observed that the Rule of Third is a viable notion in the images (especially, see Fig. 4.3.b). However, this needs more investigation. For instance, one may argue that our images are mainly taken by experts or photographers who respect such a rule of thumb.

4.5 Discussion and Future Work

In this chapter, we discussed some of the theoretical aspects of the concept of visual balance in images and artworks. We argued that this concept needs to be revisited through design mining large-scale datasets. For this account, we gathered a dataset of 120K highly liked images obtained from a professional photography website, 500px [248]. We developed a computational framework to model important or salient parts of the images with a mixture of Gaussians. Our goal was to examine what Arnheim had spatulated about the existence of stable axes and hotspots in an image. Arnheim suggested that the overlap of visual weights with these hotspots may represent a feeling of balance. Our inferred Gaussians align with Arnheim's hotspots, and confirm his structural net, specifically the center and the symmetry of the net.

At this stage, our analysis supports Arnheim's structural net. However, we believe that further investigation is necessary to understand how experts and non-experts prefer such a structure in images. Similar to some of the experiments in the recent work of McManus et al. [47], we need to study random or low liked images as well. One valid question is whether the photographers of highly liked images were aware of or even trained in some of the rules for positioning important elements in certain locations. Another question is whether by distorting the saliency structure of images we can still obtain acceptable visually balanced compositions.

Another part of our future work is to study semantics of visual balance. Because our dataset contains semantic tags (e.g. *landscape*, *architecture*, *fashion*, etc.) of the images, for future work, we aim to cluster some general visual balance templates, and to establish their linkage to the tags. This may lead to recommendations for visual layout of photos and visual design in general. For instance, in photography, it might be intuitive to layout the horizon line on one of the horizontal axes of Arnheim's net. As another example, in automatic design of magazine covers, assuming there is a human face in the cover image, we might be able to recommend a set of good candidate places to layout the face. A similar recommendation may be made for locating a brand logo in a design.

5. AUTOMATIC DESIGN OF SELF-PUBLISHED MEDIA: A CASE STUDY OF MAGAZINE COVERS

5.1 Introduction

Following progress in technology, the idea of communication has been evolving. Personal communication has been enriched by means of internet web services such as web blogs, forums, and web magazines. There are several tools that contribute to the creation of a self-publishing medium for the sake of communication over Internet. These tools support semi-automatic/automatic composition of media.

However, this idea of personal communication by means of self-publishing is multi-dimensional. One of the dimensions is to support individuals or small to medium businesses to publish by designing media. Aesthetics is an important aspect of such media. Therefore, designing aesthetically pleasing media is a conspicuous challenge that semi-automatic/automatic tools must address [258]. The main audience or customers of these tools are non-designer designers. Magcloud.com [259], for instance, supports users to publish magazines and advertisements for either web or print. Automatic design of magazines discussed in [111] is another example of such efforts. Automatic design of magazine covers is another concept that can be deployed to support current and future self-publishing work. This concept can also contribute to new approaches for the automatic design of documents and media.

Our purpose in this chapter is to address the idea of automatic design of magazine covers. We focus on magazine covers, especially because a magazine cover is one of the richest instances of a visual design. Designers spend many days conceptualizing and creating a magazine cover that will attract the audience at first glance when competing with other magazines on a newsstand [54]. If we can quantify and model the process

of designing such a medium for non-designers, we can support non-designers to produce designs that are closer to the work of professionals.

This new idea is however a multifaceted problem which includes the autonomous process of design, aesthetics and principles of design, concept in design, and cultural considerations in design. We are interested to know how professional designers think when they are in the process of creating a magazine and how they create a cover considering form and functionality. Since automatic design of magazine covers needs to be approached in an interdisciplinary manner, we have contacted professionals in schools of art, schools of journalism, and magazine editorial boards. We performed a literature review in the areas of computer science and engineering, art and graphic design, and journalism. A number of interviews with professional designers were conducted as well. We implemented what we learned in a framework of a software tool.

In this chapter, we propose a system for automatic design of magazine covers, ADoMC, that quantifies a number of concepts from art and aesthetics. Our solution to automatic design of this type of media has been shaped by input from professional designers, magazine art directors and editorial boards, and journalists. Consequently, a number of principles in design and rules in designing magazine covers are delineated. Several techniques are derived and employed in order to quantify and implement these principles and rules in the format of a software framework. At this stage, our framework divides the task of design into three main modules: layout of magazine cover elements, choice of color for masthead and cover lines, and typography of cover lines. Feedback from professional designers on our designs suggests that our results are congruent with their intuition. To the best of our knowledge ADoMC is the first system that incorporates the three main elements, layout, typography, and color for automatic visual design.

The chapter is organized as follows: Section 5.2 discusses the analysis of the contents of a cover. This section is followed by Sec. 6.3, a review of prior work in automatic design. Section 5.4 overviews our solution to automatic design. In Sec. 5.5 we discuss the layout decisions, including an adaptive template with visual balance considerations. Sec. 5.6 describes our solutions for automatic text insertion, considering the hierarchy and prior-

ity of the text contents. In Sec. 5.7 we discuss our color design algorithms. Section 6.9 presents some of the experimental results from our system. Finally, Sec.6.10 discusses the limitations of our automatic design, and provides directions for future work.

5.2 Theoretical Considerations

Through the case study of ADoMC, we take into account the key concepts of design and address the challenges that designers encounter when creating a design. The key concepts include: elements of design, principles in design, and aesthetics of design. By elements of design, we specifically mean space, type, color, and the cover photo (there are more basic elements such as line, shape, and texture that we do not aim to discuss). Principles in design, however, refer to a number of rules that a designer may apply to accomplish the task of design. Aesthetics of design deals with the beauty of designs and how good a design is, considering both the form and function of a design.

The following subsections of this section are based on the interviews with professional designers in [260], and our recollections from [44, 54, 213, 215, 216, 219, 261–272].

5.2.1 Principles of Design

Unlike visual arts where the main goals may be abstract, visual design is conceptualized and created to convey a message and communicate with audiences [6–15]. This is a key but subtle difference between visual arts and visual design. In fact, it is a critical measure of success in applied arts along with the visual appeal of the design. Therefore, unlike visual arts, there exists a set of rules so called *Principles of Design* that helps the designer to accomplish a design task. Apart from the designer’s intuition, learning and applying these rules is emphasized in design education [215, 217]. Knowing that a design message has to be conveyed at first glance, perhaps even in the first 50 milliseconds [16, 17] to make a good impression, it is critical to review some of the main principles that specifically deal with our automatic designs.

Unity

While there are different elements in design interacting with each other, it is important that they all together convey one unified message in order to comply with the design's function. A successful unity is often described by designers as clean, clear, and sophisticated. In the design of magazine covers, for instance, there might be several images and stories that the editorial board wants to depict on a cover. While the designer has to use all of these elements, he/she attempts to make one unified design.

Contrast

To direct the eyes of the reader, or to point out the important parts of a design, or to make tension and attention, contrast is applied. A more sophisticated design deliberately maintains a high contrast between the elements. For instance, for text legibility of the magazine cover, color contrast with the background, or size contrast between different texts may apply. In design, unity in variety, or a global coherence of contrasts is often advised.

Rhythm

Sense of motion and reputation of the elements may convey a form of rhythm. Apart from contrast, rhythm makes a design to be more dynamic, outstanding, and salient. Making a magazine cover design to be conspicuous is specifically important, since it has to compete with other covers on a newsstand.

Balance

Our eyes are pleased when they see a balance. Balance has different types, symmetrical and asymmetrical. Balance to designers means harmony rather than just the connotation implied by physical equilibrium (also see [247]). Visual balance however, is one of the challenging concepts in a design process. As delineated in [34], for designers visual balance in layout means "it looks right" and this can be considered as "visual weight". Some of

the well-known features that influence the notion of visual weight are light, color, shape, and texture. Although weight in terms of visual symmetry along with the vertical axis seems natural for humans (and even in animals), sometimes designers may break this type of symmetry to create tension and attention. Designers often make use of a grid (and sometimes break the grid) to create a sense of visual balance.



Fig. 5.1.: Elements of a magazine cover. (image from [273])

5.2.2 Elements of Design

The elements of design are space, text, color, and other graphics. We discuss these elements in detail in separate sub-sections in Sec. 5.4.

In order to address the concept of magazine cover design, one needs to know what the elements of a magazine cover consist of. In fact, the elements of the magazine cover customize this type of design and make it different from other media such as posters or brochures. References [54] and [262] identify the structure and elements of magazine covers through analysis of well-known magazines and interviews with professionals. As

Fig. 5.1 illustrates, a magazine cover consists of a cover image, masthead, cover lines, price, date, barcode, and some graphics called bells and whistles.

The following is a list of magazine cover elements with a concise description of the form and functionality of each element from [54].

Cover Image

Cover image is probably the most important element of the cover. Art directors, editors, and publishers have many disputes about the cover shot, since it plays such a crucial role in selling. Some research delineates that making eye contact between readers and the person on the cover contributes enormously to the sale, although sometimes designers may use an unusual shot to make the issue stand out from competitors.

Masthead

The masthead contains the title of the magazine. It has a fixed design which contributes to making it iconic, i.e. familiar to consumers. Its typeface is fixed. However, its color or perhaps its size may change. An iconic masthead gives a degree of freedom to the designer to break some of the rules about the cover's masthead.

Cover Lines

The main function of the cover lines is to absorb the reader's interest, and to convince him or her to purchase the magazine instead of a competing magazine. Cover lines also give the readers a flavor of what genre the magazine could belong to. Here, it is the form of the typeface that matters. A formal typeface indicates a different content than a casual typeface. Size and color of the text are employed to represent different blocks of information, as well as to distinguish the magazine from its competitors.

Price, Date, and Barcode

These elements present additional information to readers. Designers usually do not focus on these elements to make an interesting cover. Some magazines carry these elements in an organized structure, whereas others may put them in a random place. The price should be presented in a place that is easy to find. The date also indicates whether or not the issue is a current one. Some magazines prefer to give the actual date of the publication's date while others may prefer a sequential format to convey the longevity and prestige of the magazine. The barcode is probably the most challenging of these three elements, since the designer needs to deal with its white background. Some designers simply choose to leave it in a fixed place.

Spine

The spine makes available information related to the magazine, such as the logo (mast-head), date, and sometimes a brief itemized list of its contents.

Back Cover

The back cover may be used to make the magazine look more appealing or valuable to the reader or may be sold to advertisers.

5.3 Related Work

We believe the work of visual media's automatic design, magazine covers in particular, is a multidisciplinary task that spans several different topics, including automated layout, color design, typography, computational aesthetics, and recommendation with personalization (see the next chapter). The following sub-sections review the related work in each area and how we apply or extend prior work.

5.3.1 Automated Layout

Automated layout – how to position the textual and graphical elements in the format of documents including newspapers and magazines – has been an active area of research activity. Lok et al. [107] and Hurst et al. [109] provide surveys of automated document layout techniques. Most of the related work is based on the concept of grids in design. We also use the idea of making and breaking grids, comprehensively discussed in [264]. The idea of automated layout has recently been used in the design of photo-books [274], web-based magazines [111], and semi-automated magazine layout [112]. The main difference between prior work and our work could be the fact that we apply the layout onto a bitmap, or an image, as the background, or the canvas, rather than a white space. This requires dealing with the composition and complexity of the images as well. In fact, our definition of white space is different, since we consider the non-salient part of an image as the white space. In this paper, we will describe how we find the salient and non-salient parts of an image, and how we use them to convey visual balance in a layout.

5.3.2 Color Design

Color design is a broad concept in design that ties with several well-known color models proposed by color theorists. In computer science and engineering, researchers have deployed some of these color theories, such as color harmony and complementary colors. [40] uses Matsuda’s harmonious hue and tone templates as well as Kobayashi’s color image scale to support users for the selection of colors for product design. [41] also uses Matsuda’s harmonious templates for color harmonization in images. [149] studies color preferences provided from two color design websites Adobe Kuler [185] and COLOURLovers [148]. In our work, however, we use the well-known color theories of Itten [19] and Matsuda [40] for color design of text. In the next chapter, we apply the color semantics of Kobayashi [43] for recommendation of alternative designs to the user.

5.3.3 Typography

Typography in automatic layout is often considered as an optimization problem [109]. Knuth's work [275] is perhaps the most effective solution for typography. Recently, [276] suggests an algorithm based on Knuth's work for text fitting in arbitrary geometrical regions for digital publishing. The automatic text insertion in our typography is inspired by both [275] and [276], while dealing with text fitting in geometrical regions. Typography in design also deals with the aesthetics of typefaces and which typeface to choose for the purpose of a design [277]. Although we alternate the typeface of the textual elements in design of a magazine cover, choosing a good typeface for the purpose of a design is mentioned in the Discussion section as our future work.

5.3.4 Computational Aesthetics

Quantifying aesthetics of images is another trend of research. Birkhoff's [20] model is perhaps the first mathematical expression for aesthetics. [22] deploys Birkhoff's model for measuring aesthetics of layout of user interfaces on the screen. From another perspective, [278] develops a genetic algorithm for computing aesthetics of photo-albums. Datta et al. [28] apply a number of machine learning techniques in computational aesthetics of photos. However, machine learning techniques require defining effective features to quantify aesthetics and should be considered based on both function and form. That is why in our work aesthetics considerations are part of each module of design, e.g., aesthetics of color design or visual balance as well as the final design.

Moreover, unlike other methods, we consider high level or semantic features (e.g. "clean and clear" colors) rather than just low level features (e.g. lightness of color). From a psychological viewpoint, [26] suggests an exploratory factor analysis for this purpose and lists a number of semantic terms in aesthetics of design. However, to the best of our knowledge, deployment of high level features in automatic design has not been considered in previous work. This is, in fact, part of this thesis's contribution (See the next chapter).

5.4 Automatic Design System

This section is dedicated to how we think about the problem of automatic design from the point of view of software engineering and computer science, and how we formulate the design thought processes in a system called ADoMC (Automatic Design of Magazine Cover). We identify the main modules or processes that a typical design creation comprises. As derived from the interviews with a number of professional designers [260], we observe that for designers, the main modules in the creation of a design are perhaps layout, design of color, and typography. Hence, in order to realize the idea of our automatic design, we have devised several methods and mathematical expressions to quantify and perform each module automatically. In this section, we first present a schematic overview to our software framework in Sec. 5.4.1. We then discuss in detail each phase of the schematic overview in specific sections. In Sec. 5.5 we talk about the concept of layout which in turn includes the visual balance considerations. We talk about how we have customized this concept in our automatic design. Section 5.6 dedicates a discussion to the concept of typography and the way that we have deployed it in our automatic design to make more aesthetically compelling designs. Section 5.7 talks about the concept of color design and how we formulate some of the well-known color theories in the framework of our automatic design.



Fig. 5.2.: A schematic view of ADoMC. Image reproduced from [110].

5.4.1 Software Framework Overview

As Fig. 5.2 shows, our framework includes the three main modules of the design in the format of three phases to accomplish the task of automatic design. Note that in our software system, typography and color design work independent of each other. However, any change

in the user inputs makes the system start all over again and design from the beginning. This decision is very important: any change in a design may lead to a whole new design [260]. That is part of the reason that we do not allow the user to apply any modification to our design. In fact, we learned that if the user does not know how to design, he/ she can easily make big uglies [260]. Also, note that our system starts the design from a well-composed image as input.

The system first takes a well-composed image as input. Here, we define a well-composed image as an image which obeys the rule-of-thirds [279]. This rule suggests that humans like to see important parts of an image to be placed on the two vertical and horizontal imaginary lines on the image. Another feature of a well-composed input image is to have the aspect ratio of 8.5in width to 11in height. This is a standard size for magazine covers. Note that resizing or cropping an image in order to make it fit for a specific purpose is a whole separate topic. There are some effort has already been made in the field of image processing and computer graphics for photo cropping. However, to confine the boundaries of our problem, we assume that the cropping and resizing according to the cover image has been done and a well-composed image will be entered by the user to the system.

Given a well-composed image, the system makes decisions the layout of elements. In order to have an aesthetically appealing layout, one needs to consider visual balance. In order to quantify visual balance, the system employs a visual saliency algorithm ([164] and [279]) to find important parts of the cover image. Here, important or salient parts are the most conspicuous regions of an image based on prediction of human eye fixations. In this fashion, the framework can designate a number of regions as candidates for inclusion of other elements of the cover, such as cover lines. Moreover, we compute the visual balance of the cover image to narrow down our choices of the candidate regions. Section 5.5 talks about the layout and visual balance phase in more detail.

Insertion of text in the candidate regions, obtained in the layout and visual balance phase, is part of typography. Section 5.6 formulates the task of automatic text insertion in the geometric candidate regions as a number of optimization problems. This text insertion follows some rules and patterns recognized by analyzing the work of designers.

Design of color for the masthead, headline, byline and other cover lines of the magazine is the other phase of the system. Our system makes a number of decisions based on the color distribution of a given cover image and then designs a color palette for the magazine cover. Section 5.7 collects two well-known color theories and devises a number of mathematical expressions to quantify these color theories to deploy in our automatic design. This section also identifies another important challenge in design of color for texts, which is text legibility. Hence, a color modification process is devised and applied after design of the text color palette in Sec. 5.7.3.

5.5 Layout

As noted in [107], layout can be considered as a process that tries to solve the problem of positioning the visual elements while considering or determining their sizes. However, creating aesthetically pleasing layouts ties with visual balance considerations in design and effective usage of the white space. In design, the form of the layout should follow the functionality of the design. For instance, the layout of the magazine cover should be in a form that directs the eyes of the audience through the design. In the layout process, the designer comes up with a structure to position and scale the design elements interacting together yet together representing a unified look.

The layout of a magazine cover also has to have an appealing form to attract the reader at first glance while it is on the stand of magazines [260] and [54]. In order to meet both required function and appealing form, designers apply some imaginary grids on the sheet of the arrange the visual elements of the design through these grids. Designers may occasionally break the grids as well [264]. Making and breaking grids is a concept in design which has been influenced by modernism and later postmodernism art movements [280].

Our space is the entire cover image. We need to define the white space and the hierarchy while considering balance and contrast principles discussed in Sec. 5.2.

5.5.1 Defining Space

As function of the magazine cover requires it, some of the elements on the cover should be laid out in fixed positions. These positions are mainly justified based on the priorities and the hierarchies the editorial board of a magazine assigns to different elements. More specifically, the layout of the magazine cover should give a visual hierarchy and a sense of priority to the stories that have been decided to be presented on the cover, and lead the reader's eyes through the stories. Among the elements of a magazine cover, the masthead is always positioned at the top of the cover. It also has to be as large as possible, to blast the magazine's name. Based on the rules in Sec 5.2, we have derived a general template for our automatic magazine cover design illustrated in Fig. 5.3.

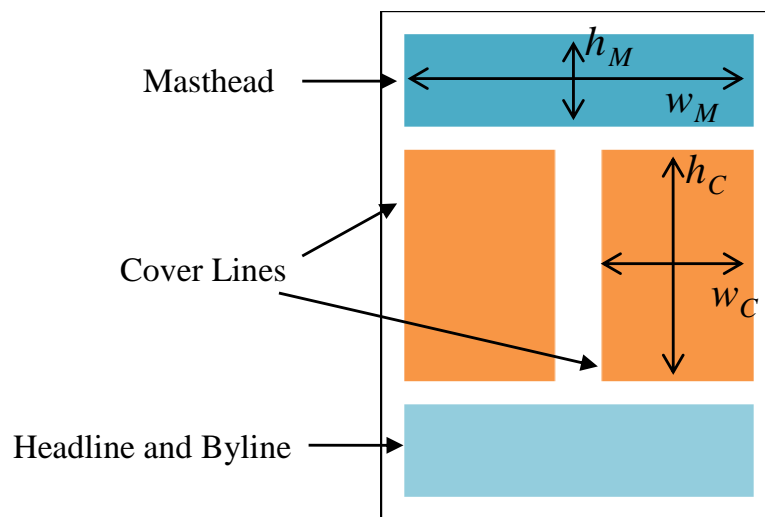


Fig. 5.3.: A general template for magazine covers.

This template in fact is based on the concept of grids in design. In this fashion we can constrain the problem of layout to a more specific form which is also aligned with the functionality of the magazine cover. The alignments and margins help us to design neatly. Also note that base on the content of a given cover image, we may break some of the grids. This idea provides a general template and makes the produced designs look more customized to their background images.

More specifically, this template can vary based on the visual balance considerations which are unique and specific for each design and mainly influenced by the cover image of a magazine. The only caveat attached to this current template is that we consider a fixed position for the headline and byline. This decision is made to present the important stories of the magazine in an isolated part, and yet distinguishable from the masthead. In other words, as learned in the interviews with the professionals, we do not want the headline and byline to be in a place that visually competes with the masthead on the priority criterion.

5.5.2 Visual Balance Considerations

Having devised a general template for our magazine covers, we need to consider the visual balance concept to utilize its predefined grids. As discussed in Sec 5.2 in visual balance, our goal is to optimize the usage of the white space to represent a sense of visual weight, and “looking right”. These terms, the work of [22] in measuring aesthetics of layout, and the grid-based approach in layout [264], form the quantification approach of the visual balance of magazine cover layout in ADoMC, illustrated in Fig. 5.4.

For visual balance considerations, ADoMC first finds the white spaces of the cover image. White space on the background image means the regions of the image which are less busy relative to the other regions of the image. That is, ADoMC finds the non-salient segments of the image by deployment of the Graph-Based Visual Saliency algorithm (GBVS) [164]. ADoMC then fuses the binary GBVS mask with the general magazine cover template. The result yields the candidate regions for holding the textual elements on the cover. Finally, ADoMC applies grids on the image and assigns different weights to different cells of the grids to compute which candidate region should hold the textual elements. In this fashion, ADoMC conveys the visual balance in the layout of magazine covers. However, there are more details that we need to address to be able to find which candidate region is the best among all the candidates. In other words, we want to know whether we insert the cover lines to either the left or the right side of the cover; or in some occasions, distribute the cover lines to both sides of the cover.

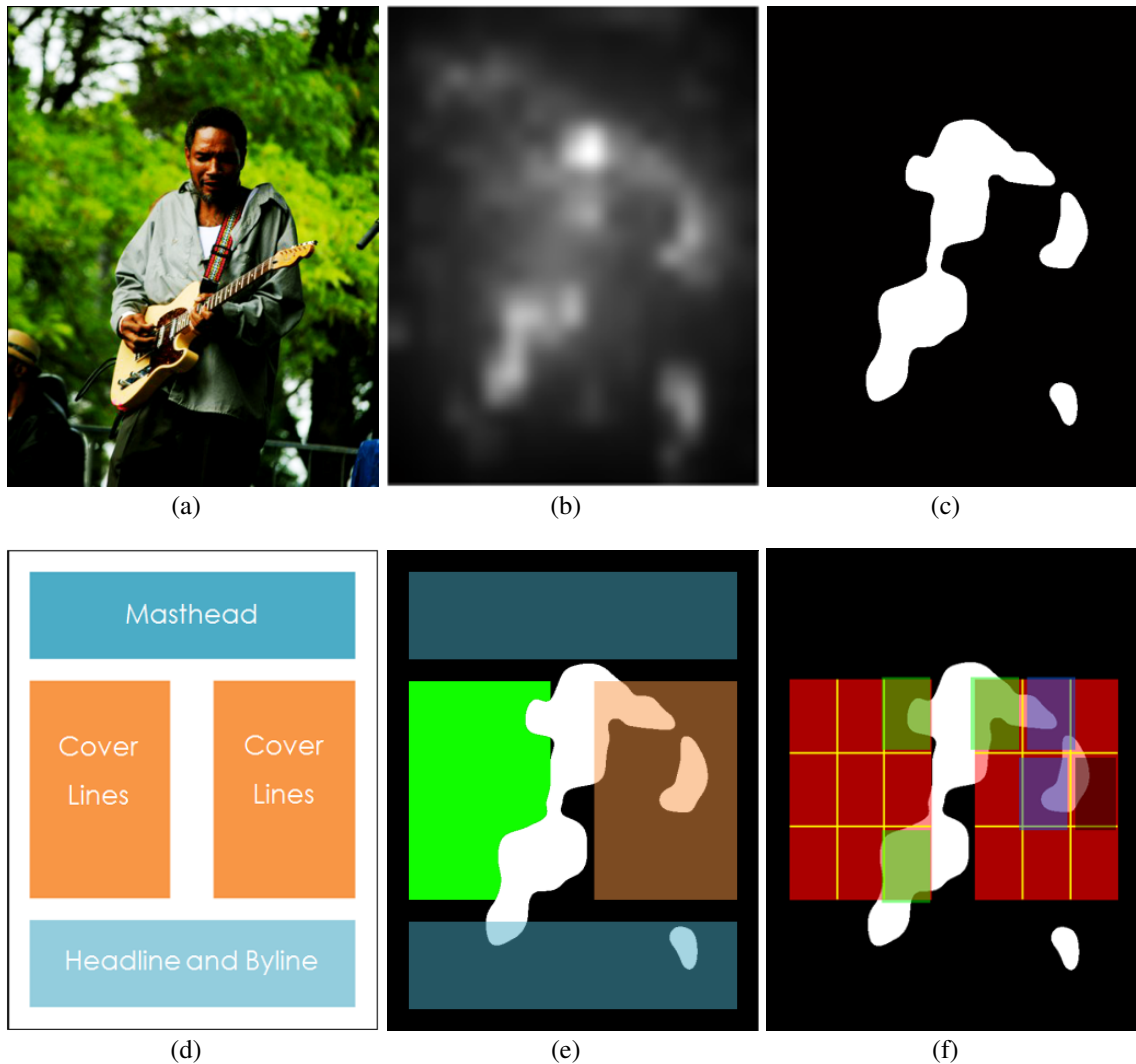


Fig. 5.4.: Layout decisions and visual balance considerations. a) An image. b) GBVS [164] saliency map. c) Binarized mask from (b). d) A general template for magazine covers. e) Fusion of (c) and (d). ADoMC keeps the green region and discards the brown region, since the green region looks empty and is a good candidate to hold the cover lines. This decision is made by computing the visual balance as shown in (f). f) Computation of visual balance: different weights are assigned to the cells of the yellow grids on the left and right sides of the cover image: the darker the cells, the heavier the weights. Summing up the product of each cell's weight and the number of salient pixels (here in white) gives us the weights of the left and right grids (see Eq. 5.2). The side with a lighter weight will be considered as an empty space to hold more elements, such as the texts of the cover lines (see Eq. 5.3). Photos courtesy of Amir Garachorlou.

As Fig. 5.4.b shows, GBVS gives us a mask which contains values between 0 and 1. In order to obtain a binary mask shown in Fig. 5.4.c, we threshold the GBVS mask. This threshold depends on the cover image. We use the Newton-Raphson method to obtain $\frac{1}{9}$ of the total area of the salient parts. We choose the fraction $\frac{1}{9}$ based on the intuition that if we consider the rule-of-thirds, then $\frac{1}{9}$ of an image is the most conspicuous or important part of it.

Having obtained the GBVS binary mask (Fig. 5.4.c) and fusing the general cover template (Fig. 5.3) to it, we obtain Fig. 5.4.d. The green region in Fig. 5.4.d is considered as a candidate for holding cover lines. However, the question is how to identify when the green region wins over the other side (the dark orange region). Figure 5.4.f illustrates how we apply grids (yellow grids in Fig. 5.4.f), and assign weights to the both left and right sides (green and dark orange). In this fashion we find which side is the better choice as the candidate region for inclusion of the cover lines.

As illustrated in Fig. 5.4.f, we consider two grids on the left and the right side of the cover image. Then, different weights are assigned to the cells of the yellow grids: the darker cells (refer to Fig. 5.4.f) are assigned with the heavier weights. Summation of multiplication of each cell's weight to the number of the salient pixels (here in white) that it contains gives us the total weight of each the left and the right grids. This represents the fact that while our eyes move away from the center of the magazine cover, we are ready to see more text. This idea is based on consulting the work of professionals. In fact, in the most of magazine covers, designers try to wrap around the salient part(s) of the cover image with the cover lines. That is why we do not want to put the cover lines on a side with more salient part close to the borders of the cover. Equation 5.2 shows how the weights for the left side and the right side (red regions in Fig. 5.4.f) are computed. Based on these two weights we choose the final candidate region for the inclusion of the cover lines, according to Eq. 5.3.

We denote w^L and w^R as the obtained weight for left side and right side, respectively. Also, g_{ij}^L and g_{ij}^R denote each cell in the left grid and the right grid (yellow grids in Fig. 5.4.f), respectively. g_{ij}^L (and similarly, g_{ij}^R) is obtained from the following equation:

$$g_{ij}^L = \begin{cases} n & \text{if } n \geq T \\ 0 & \text{otherwise,} \end{cases} \quad (5.1)$$

where n is the number of white pixels in the ij -th cell. Here, based on our experiments, $T = 800$ is satisfactory.

Given g_{ij}^L , we compute w^L (similarly, w^R) as:

$$w^L = \frac{1}{6} \sum_{i=1}^3 \sum_{j=1}^3 j \cdot g_{ij}^L. \quad (5.2)$$

The final candidate region for inclusion of the cover lines is obtained by:

$$c = \begin{cases} 1 & \text{if } w^L < w^R \\ 2 & \text{if } w^L > w^R \\ 3 & \text{if } w^L = w^R \end{cases} \quad (5.3)$$

Here we index each condition with 1, 2, or 3, indicating left side, right side, or both left and right sides, respectively. Note that if the third condition happens for c , we use both sides for the automatic text insertion process. The next step is to wrap the cover lines in the candidate region, c . This process is discussed in Sec. 5.6.3 as an automatic text insertion process.

5.6 Typography

Where and how to put text, including the masthead, headline, byline, and other cover lines, is an important decision in design of magazine covers [54]. In our problem scope, by typography we mean the decisions that a designer makes about using different typefaces for

different design purposes, the text size of the cover lines, priorities of the cover lines, and the problem of legibility of the text. According to professionals, the family of *Helvetica* is a good choice for a typical magazine cover. Better choices, however, relate to the genre of magazine and will be considered in our future work.

Similar to the preceding discussion for layout, we need to convey the priority and the hierarchy of the cover lines on the cover. Therefore, ADoMC assigns different sizes to the cover lines and lays them out based on their priority (specified by the user) starting from the top of the cover. Given a candidate region from the layout process, ADoMC inserts the text by applying a special form of indentation. This indentation is a pattern derived from well-known magazine designs. That is, designers usually give an indentation to the texts to align them with the boundaries of the salient segments of the cover image. Therefore, they can use the white space in an effective way. This technique can be formulated as how to insert or fit a text in a geometrical region such that its indentation is most faithful to the shape of the region. ADoMC solves this problem as an optimization problem in which the maximum fidelity to the shape of the geometrical region and the maximum size of the text is desirable (see Sec. 5.6.3). Figure 5.5 illustrates the results of this solution for different numbers of cover lines with different words.

5.6.1 Masthead

As Fig. 5.3 illustrates, the masthead is always positioned at the top of the cover. It also has to be as large as possible, to blast the magazine's name.

First, we define the cost function $d_M(w, h)$ as a function of width w and height h of a given masthead M . Note that M can be either a predesigned logo or a string of text, entered by the user. In the latter case, width and height of the surrounding box (bounding box) of the string are the parameters of the cost function:

$$d_M(w, h) = d_w(w) + d_h(h), \quad (5.4)$$

where $d_w(w) = w_M - w$ and $d_h(h) = h_M - h$, with w_M and h_M as our predefined values for maximum width and height of the masthead in the general template, illustrated in Figure. 5.3.

Then, we define $F_M(w, h)$ as the optimization function which obtains the optimal values of width and height of the masthead. We denote (w_0^M, h_0^M) as the optimal masthead signature for a given masthead M , such that

$$(w_0^M, h_0^M) = \underset{w, h}{\operatorname{argmin}}(d_M(w, h)), \quad (5.5)$$

for $w_m \leq w \leq w_M + \Delta w$ and $h_m \leq h \leq h_M$. Here w_m and h_m stand for minimum of width and height, respectively. Also, Δw and Δh stand for some small numbers (in inches) as the tolerances for cases where a small change in the sizes of the predefined values for width and height may obtain a better result.

5.6.2 Headline and Byline

Similar to the way that we find the optimal size for the masthead, we can find the optimal sizes of the typeface for the headline and the byline texts, positioned at the bottom of our automatic generated covers (see Fig.5.3).

We denote $F(H)$ and $F(B)$ as the optimization functions for the case of the headline and the byline text, respectively. The only difference with the case of the masthead is that since we have to convey hierarchy and priority of the stories, we choose a smaller size of typeface for the lower priority story. Hence the maximum width and height for the headline is one unit less than the obtained values for the corresponding parameters in the masthead case from the masthead signature. Headline signature is then denoted by (w_0^H, h_0^H) , and computed as:

$$(w_0^H, h_0^H) = \underset{w, h}{\operatorname{argmin}}(d_H(w, h)), \quad (5.6)$$

for $w_m^H \leq w \leq w_M + \Delta w$ and $h_m^H \leq h \leq (h_0^M - 1) + \Delta h^H$.

Based on the same rationale for hierarchy depiction by using different text sizes, the maximum height of the byline should start off with at least one unit smaller than the obtained value for the height of the headline from the headline signature.

Finally, in the insertion of cover lines, the maximum height of the cover lines should be at least one unit smaller than the obtained value for the height of the byline from the byline signature.

5.6.3 Cover Lines

For insertion of cover lines however, we want to make a special form. This form is inspired by looking through the work of designers. In fact, although we consider the candidate region from Sec. 5.5.2 (see Fig. 5.4) to put the cover lines, we want to make the design more compelling and closer to the work of professionals. By looking at different magazine covers we have found a general pattern of how to make text indentation for cover lines. Indentations are usually in a form that imitate boundaries of the salient part(s) of the cover image. This design decision makes a dynamic interaction between the design elements and leads to a more unified design.

To illustrate this pattern and our idea for reproducing such patterns in our automatic designs, consider Fig. 5.5.a (obtained in the same way described in Sec. 5.5.2) which contains a green region. The borders (red curve) of this region can determine how the indentation of cover lines in this region should look like. Figure 5.5.b and Fig. 5.5.c illustrates the result of our solution for this problem.

Therefore, our problem of how to give this kind of indentation to the cover lines can be defined as the following: how to insert the text such that its indentation has the maximum fidelity to the shape of the green part and simultaneously make sure that its size is the largest in an acceptable size range for cover lines. Figure 5.6.a and Fig. 5.6.b illustrate that with a smaller size of typeface of the text—the blue boxes represent the bounding box of each line of the text—we may obtain a higher fidelity to the geometric region—the orange

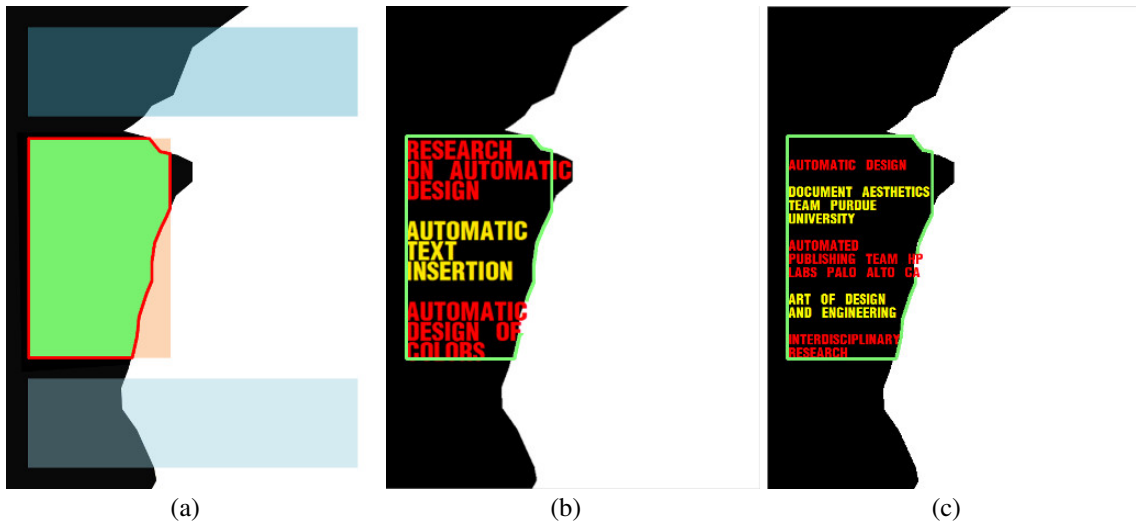


Fig. 5.5.: Examples of automatic cover lines text insertion on a magazine cover. a) The green region is a candidate to hold cover lines. We want to insert the text of the cover lines in such a fashion that the text indentation follows the shape of the green region and the size of the text is as large as possible. b) Result of our automatic text insertion in the green region in (a). c) Another result of our automatic text insertion in the green region in (a). Note that the number of words and different words in the cover lines yield different results.

shape. On the other hand, the larger size of the text is preferred to obtain a better legibility of the text. Therefore, we define this problem as an optimization problem where maximum fidelity to the shape of the candidate region (the orange region in Fig. 5.6) and maximum size of the text is desired. As Fig. 5.6.c and Fig. 5.6.d show we define a distance penalty or a cost function as the length of the red arrows and we minimize this cost function while trying to keep the size of the text as large as possible.

We define the optimal text insertion parameters' values by this signature:

$$(s^*, y^*, \mathbf{d}^*, h^*) = \underset{s}{\operatorname{argmax}} (\underset{y, \mathbf{d}, h}{\operatorname{argmin}} (F_{CL}(s, y, \mathbf{d}, h))), \quad (5.7)$$

where

$$F_{CL}(s, y, \mathbf{d}, h) = \sum_{k=1}^K d_k^{(s)} + y^{(s)} + h^{(s)}, \quad (5.8)$$

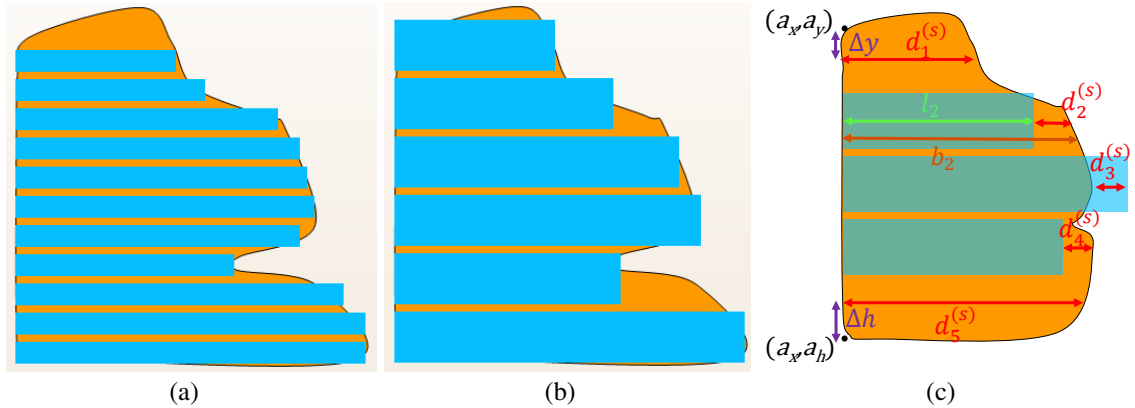


Fig. 5.6.: Insertion of cover lines on a magazine cover. The blue boxes represent the bounding box of each line of the text and the orange geometric regions represent the candidate regions for inclusion of the text. a) A smaller size of the text may result in a higher fidelity to the candidate region of the text whereas b) shows that larger size of text is also desired. c) shows the distance penalties (in red and violet) and other variables. These distance penalties represented by red and violet arrows need to be minimized.

for $s \in [18, 36]$, where $|y^{(s)} - a_x| \leq \Delta y$, $|h^{(s)} - a_h| \leq \Delta h$, and \mathbf{d} , contains the distances $d_k^{(s)}$ for each line k defined by $|l_k - b_k|$. See Fig. 5.6.c for the variables.

5.7 Design of Color

Design of a color palette —or which colors to choose to create a good design— is part of a designer’s work and concern. There are resources [148, 185, 281–283] that suggest which colors go with which. In general, designers try to make a contrast or a linkage between colors [43]. However, our problem is how to quantify these concepts. One of the most significant works in the area of quantifying aesthetics of colors is Itten’s color harmony concept [19]. Itten defines seven kinds of color contrast: contrast of hue, light-dark contrast, cold-warm contrast, complementary contrast, simultaneous contrast, contrast of saturation, and contrast of extension. From these contrasts, only complementary contrast is used in our work. Itten’s complementary contrast concept is based on physiological

laws of afterimage and simultaneity. These laws suggest that our eyes are satisfied when they can perceive a balance in colors —or in Itten’s words “complementary colors”. Itten then defines a hue wheel as shown in Fig. 5.7.a, which presents complementary colors in opposite directions. We apply the concept of complementary contrast in our automatic framework to choose the color of the masthead.

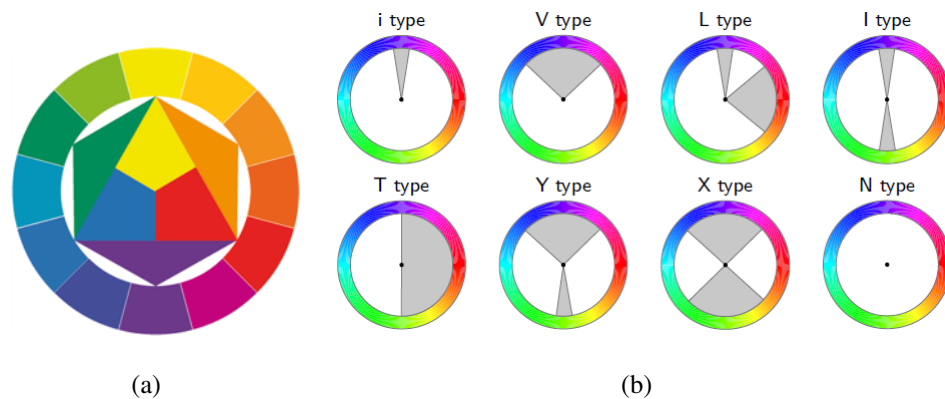


Fig. 5.7.: Elements of color design. a) Itten’s hue wheel. Note that complementary colors are in opposite directions: for instance, yellow and violet, or orange and blue. b) Matsuda’s harmonious templates. This figure is from [41]. Hues which are in shaded sectors are considered to be in harmony. The angles of the sectors are described in [40] and [41].

Having chosen a color for the masthead, now we want to find a set of colors for cover lines. For cover lines, we can use the idea of the *similar hues* and the *same hues* discussed in [284]. Some color resources call these ideas analogous colors [185] or light-dark contrast [39]. Similar hues are color hues which are adjacent colors on the hue wheel. Same hues are colors in the same hue but with different levels of shade [284], or more precisely, different values of S (saturation) and V (value) in HSV color space. A more comprehensive definition of these concepts is suggested by Mastudas harmonious templates [40, 41]. Figure 5.7.b illustrates seven different hue templates and a template for gray-scale images (N type). In each template, the hues that are in shaded sector(s) are considered to be harmonious. Later in Sec. 3.5.2 we talk about how we use Matsudas harmonious hue templates to design the colors of cover lines.

In order to work on the hue wheel of an image, we first convert the color space of the image from sRGB to HSV —Hue, Saturation, Value. The reason is that HSV has been defined to support an intuitive color space for designers. One of the reasons that HSV is intuitive is perhaps because of the fact that it supports the realization of the subtractive colors concept versus the additive colors concept. In the real world, as opposed to the world of computers, what we can see as far as colors and color mixtures are considered they are subtractive colors. In fact, the color that we perceive is a subtracted component of the light that has been illuminated to an object surface. In contrast, the colors that computers presents to our eyes are the results of an additive color system. In means that computers add several light components to represent a color corresponding to a real color. This additive color system is implemented in the sRGB color space. In order to transform between the real colors —the colors that designers use in reality— and the computer generated colors, we frequently switch to HSV and sRGB. We then work on the hue channel which is defined from 0 to 360 (degrees), where 0 corresponds to red. This means that we keep the two other components of the color, S and V as their maximum amounts ($S = 100$, $V = 100$). In this fashion we just work with the colors located on the border of the circle of the HSV cone.

5.7.1 Masthead Color

For color of the masthead, we use the idea of the complementary colors suggested in Itten's hue wheel. As Fig. 5.7.a illustrates, Itten has chosen ten colors as representative of all the colors. However, taking into account the capabilities of computers and printers, we can support more than ten distinct colors for our automatic design of color. Therefore, we have upsampled Itten's hue wheel to 360 hues. Another fact is that Itten calls these ten colors ten hues by which he means ten pure or highly saturated colors. Correspondingly, we should define our mathematical framework in a color space which can be an appropriate representative for the hues that Itten meant.

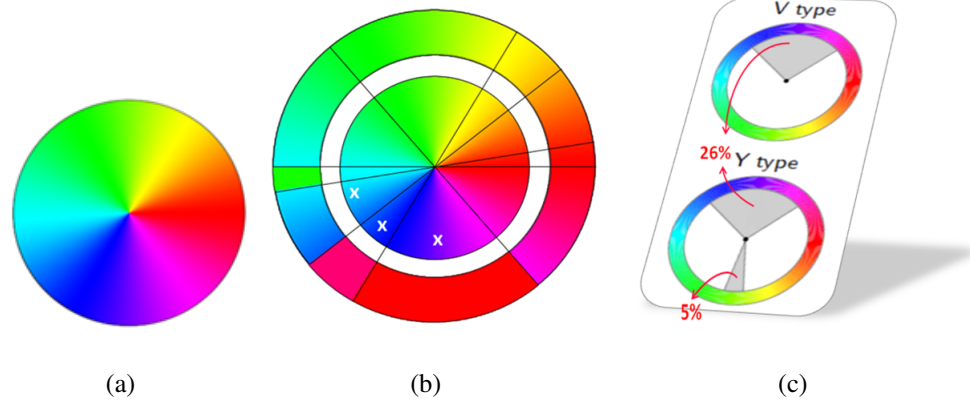


Fig. 5.8.: Our choice of hue wheel for computing the complementary color. a) A continuous pure hue wheel. This means that saturation and value are equal to the highest possible amounts. b) The inner wheel shows the hue wheel. But we discard the sectors marked with “×”. As a result, we use the outer wheel to return the computed complementary (see Eq. 5.10). c) Our design of color palette will have a triple of hues, all in the shaded sectors of V-type and Y-type templates of Matsuda.

In order to construct Itten’s hue wheel, we first transfer from sRGB color space to HSV color space. We then consider the Hue component of the color and assign the highest amounts to Saturation and Value components. In this fashion we obtain a continuous hue wheel which later will be quantized or downsampled to 360 hues. Figure 5.8.a illustrates the result of this decision.

As one can notice, the result is not a complete equivalence of Itten’s hue wheel in respect to the opposite hues. Therefore we slightly change the hue wheel to obtain more artistically intuitive results. These changes are depicted in Fig. 5.8.b. More precisely, we define the membership function for the complementary color as following:

We denote the hue of the masthead color as H_M which is the complementary hue of the most repeated hue in a given cover image, and obtained from the function $compl(h)$:

$$H_M = compl(h_0). \quad (5.9)$$

According to Fig. 5.8.b, we define function $compl(h)$ as:

$$compl(h) = \begin{cases} 115, & \text{if } h \in [180, 190] \\ 333, & \text{if } h \in [220, 239] \\ 0, & \text{if } h \in [240, 309] \\ (h + 180) \bmod 360, & \text{otherwise.} \end{cases} \quad (5.10)$$

Where h_0 in 5.9 is the most repeated hue, obtained from:

$$h_0 = \mathit{archist}_M(\mathit{max}_{hue}). \quad (5.11)$$

Hence, the color of the masthead, denoted by $color_M$ is obtained by:

$$color_M = RGB(H_M, S = 100, V = 100). \quad (5.12)$$

5.7.2 Cover lines Colors

Given the hue of the masthead H_M , we obtain the color of cover lines, namely headline color, $color_{HL}$ and byline color, $color_{BL}$ from 5.13 and 5.14, respectively. In this fashion, we obtain our cover line colors based on the idea of similar hues in V type and Y type templates of Matsudas harmonious hue templates.

$$color_{HL} = RGB(H_{HL}, S = 100, V = 100), \quad (5.13)$$

for $H_{HL} = (H_M + 15) \bmod 360$.

$$color_{BL} = RGB(H_{BL}, S = 100, V = 100). \quad (5.14)$$

for $H_{BL} = (H_{HL} + 15) \bmod 360$.

So far, we have designed the color of the texts based on the hue wheels. Part of the motivation for using hue wheel is that, hue colors are vivid, and suitable for color palette of

the magazine covers. In fact, using vivid colors is a form which follows the functionality of the color of the texts in the magazine covers. However, we occasionally apply the Lightness Contrast to enhance the legibility of the texts while keeping the color design palette aesthetically compelling. The next section describes this idea in more details.

5.7.3 Color Aesthetics and Text Legibility

In the previous section, we described how we have quantified and computed the concept of color harmony for our automatic design of color. However, in practice there are some challenges that we need to address. The main challenge is perhaps that while the color palette design has to be aesthetically pleasing, it also has to be simple. Therefore we should limit the palette to a pair or a triple of colors. This is an enormous constraint that we put: sometimes these two or three colors are not very legible on the cover. So we have to come up with a better color, subject to the legibility criterion. One way of solving this problem is to expand our use of color and include the variation of Saturation and Value in the design of color. Matsuda also defines his tone templates, where he presents the tones that are considered to be harmonious. Figure 5.9 illustrates Matsuda's harmonious tone templates.

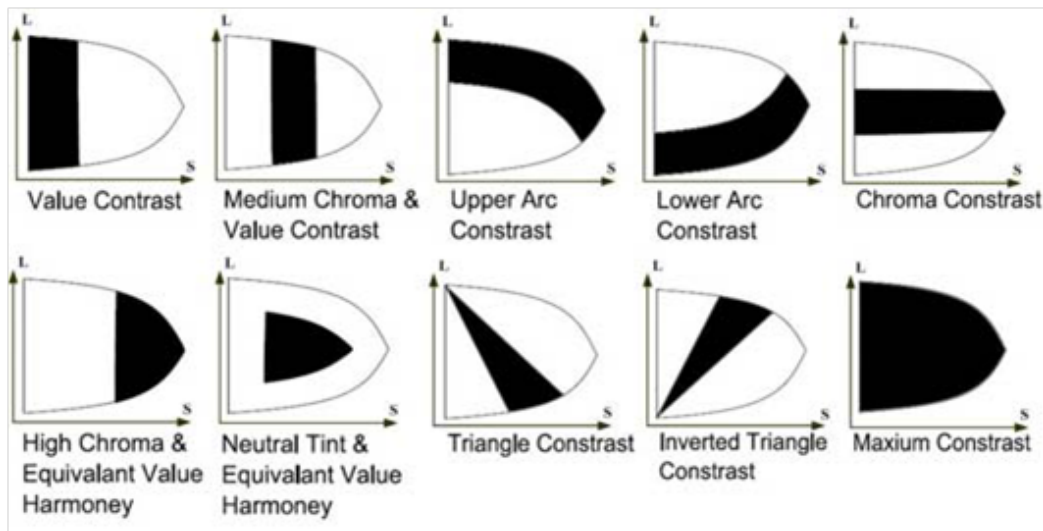


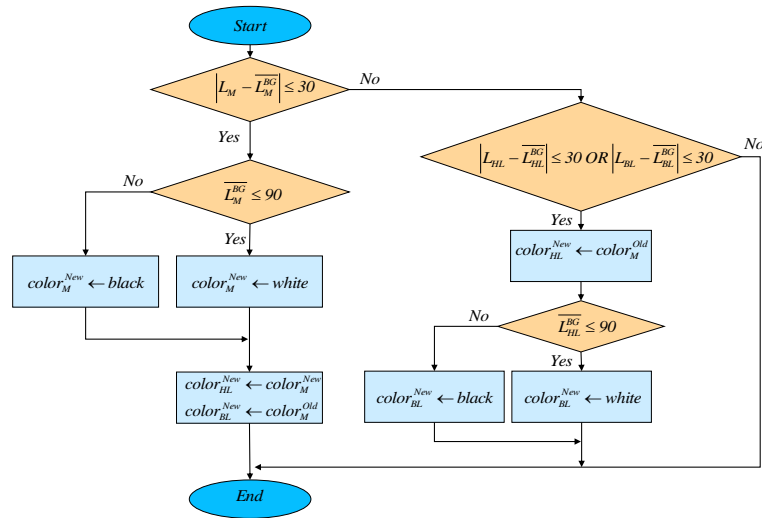
Fig. 5.9.: Matsuda's harmonious tone templates. a) Tone templates. This figure is from [41]. Tones which are in the shaded sectors are considered to be in harmony. The measures of the sectors are described in [41].

Similar to what we have done for hues, we need to find a corresponding framework for tones in computers. Matusda's tone coordinate is composed of Chroma versus Value. Here, Chroma and Value are representative of Munsell's Chroma and Value in Munsell's color system. In fact, Matsuda designed his tone templates in the same shape of Munsell's Chroma-Value coordinates. Matsuda then marked the regions that he found in harmony. Figure 5.9.a shows that marked regions in black, indicating any two points in a black region are considered harmonious. In our design, we use the Upper Arc Contrast and Lower Arc Contrast tone templates. More specifically, we choose two points in these two black regions: the tip and the tail of the arc. This gives us a combination of a hue (i.e. a vivid color, similar to what we designed in previous sections) and of either a white color or a black color. The justification of this decision comes from the fact that in our contemporary design, professionals tend to use white (in majority) and black (in minority) in their designs.

In order to choose between white and black, we first consider the local background lightness of the text. If the lightness of the local background is not higher than a threshold we use white, otherwise we use black. Note that in order to compute the lightness of

the colors we transfer the colors from sRGB color space to color space, since this color space is perceptually uniform and an appropriate system for representation of lightness and noticeable contrast to human eyes. In our framework, we evaluate the lightness of the colors and the local backgrounds by just considering one component.

One way to solve this problem is to expand our use of color to include the variation of Saturation and Value in the color's design. Going back to the geometric structures, we observe that Matsuda has defined tone templates, as seen in Fig. 5.9. These tone coordinates are composed of Chroma versus Value. Here, Chroma and Value are representative of Chroma and Value in Munsell's color system. The colors seen in the shaded regions are considered to be harmonious. In our design, we use the Upper Arc Contrast and Lower Arc Contrast tone templates (the middle column in Fig. 5.9. More specifically, we choose two points in these two black regions: the tip and the tail of the arc. The reason for this decision is that in our contemporary design, as we have stated above, professionals tend to use white and black for text colors. Similarly, in our designs we use a pure or highly saturated color with white or black. In order to choose between white and black, we first consider the local background lightness of the text. If the lightness of the local background is less than a threshold, we use white. Otherwise, we use black. This process is implemented in the flowchart found in Fig. 5.10.a. It is important to note that the comparisons in Fig. 5.10.a are computed in $CIE L^*a^*b^*$ color space because this color space is perceptually uniform, and is a useful system for representation of lightness and just noticeable contrast to the human viewer. Also, the numerical values of the thresholds in Fig. 5.10.a are based on psychophysical experiments, some described in [285].



(a)

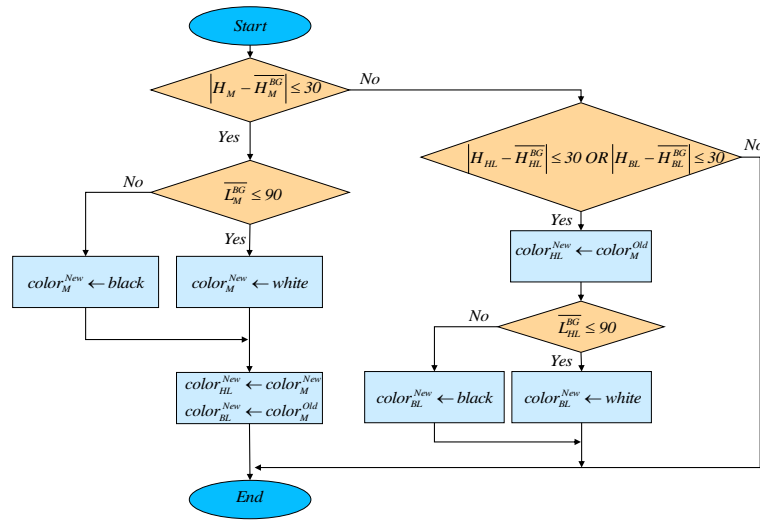


(b)



(c)

Fig. 5.10.: Text legibility solution and results using hue modification. a) Flowchart representing color modifications for text when the lightness (in $CIE L^*a^*b^*$) of the text and its local background are too close. In this flowchart, the lightness of the masthead is denoted by L_M , and the mean lightness of the masthead's local background by L_M^{BG} . Also, HL and BL indicate headline and byline, respectively. The numerical thresholds are based on psychophysical experiments, some described in [285]. b) and c) Before and after applying the flowchart in (a), respectively. Photos courtesy of Andy Jessop.



(a)



(b)



(c)

Fig. 5.11.: Text legibility solution using lightness modification. a) Flowchart representing color modifications for text when the hue (in $CIEL^*a^*b^*$) of the text and its local background are too close. In this flowchart, the hue of masthead is denoted by H_M , and the mean hue of the masthead's local background by H_M^{BG} . Also, HL and BL indicate headline and byline, respectively. The numerical thresholds are based on psychophysical experiments, some described in [285]. b) and c) Before and after applying the flowchart in (a). Photos courtesy of Andy Jessop.

Another approach to solving the text legibility problem is to apply shadow effects to the text. This approach is often taken by designers to mitigate the problem in an aesthetically pleasing fashion. Similarly, in cases where illegible text colors occur, we provide an alternative design for the user that includes text shadow effects. Fig. 5.12 illustrates alternative designs generated by our system.



(a)

(b)

Fig. 5.12.: Text legibility solution using shadow effect. An alternative solution to improve text legibility by applying shadow effects to the text, a) before, and b) after modifications, respectively. Photos courtesy of Andy Jessop.

5.8 Experimental Results

We have implemented the the automatic framework discussed in the previous sections as a software application. We have named this system ADoMC: Automatic Design of Magazine Covers. As specified earlier, the users of our system are non-designer designers; from the range of individuals, small to medium businesses. Therefore, ADoMC has a simple graphical user interface (Fig. 5.13), because we want to keep the creation of the

design for our framework without the possibility of modification from the user. The reason is that we think if the user does not know anything about design, he/she can easily make “big uglies” (noted in [260]). Figure 5.13 illustrates a screenshot of the graphical user interface of our system. Our system does all the tasks automatically with just a click. It takes a few minutes for the system to accomplish the entire design. Our code is in Matlab and our experiments have been performed on a 64-bit PC machine with a 2-core 3GHz Intel processor and 4GB of RAM. This system will be implemented in Java as an internet service.

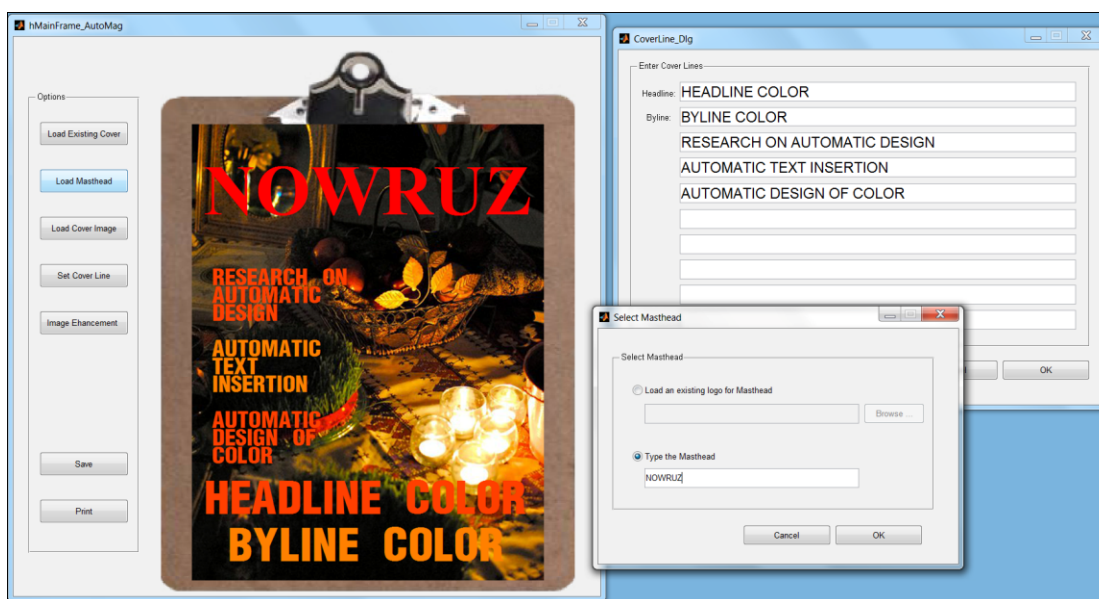
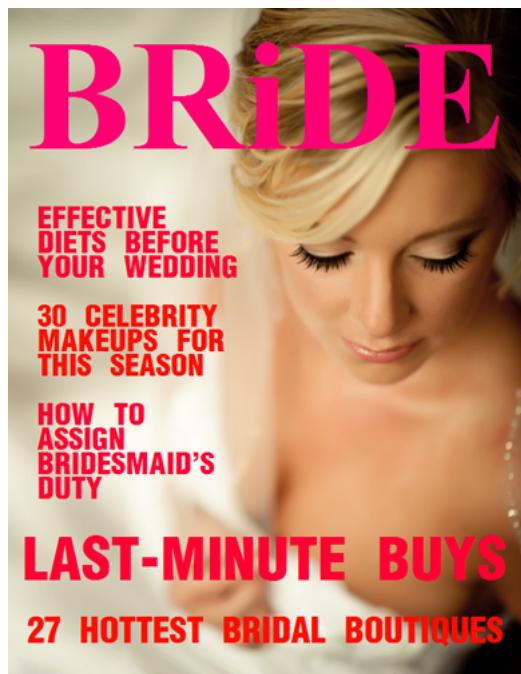


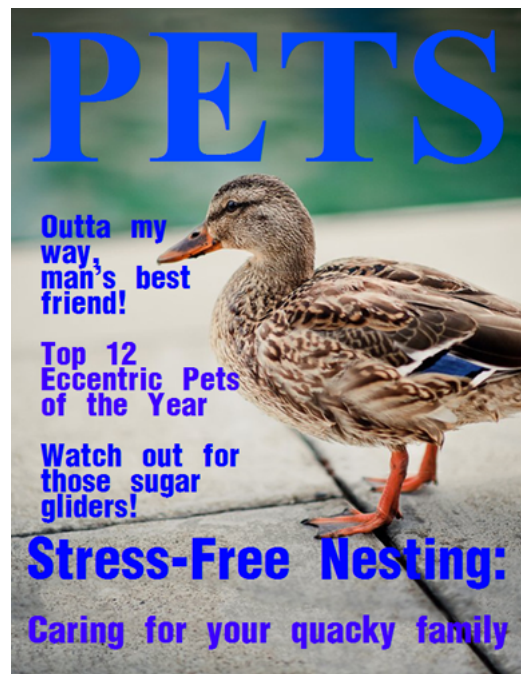
Fig. 5.13.: A screenshot of the graphical user interface of ADoMC (Automatic Design of Magazine Covers) system.

5.8.1 Results

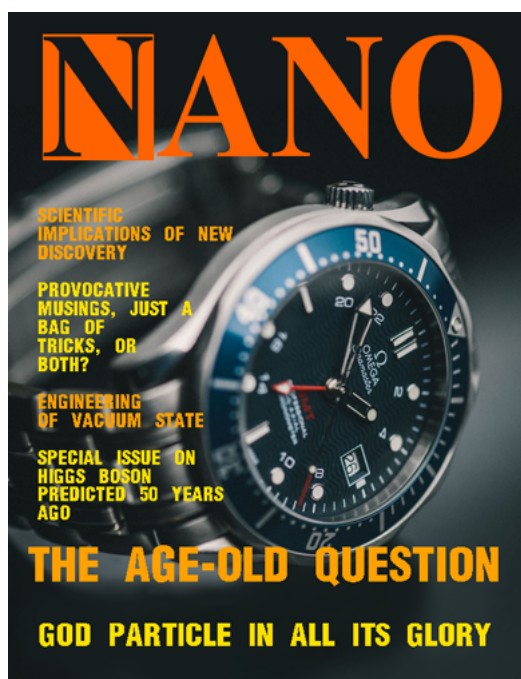
In this section, we present four more automatically designed magazine cover samples by ADoMC. Figures 5.14 (a,b, and c) illustrate cases where there was no need to perform a color modification. Therefore, the colors have been designed based on the hue templates. Figure 5.14.d however, illustrates a case in which the color modification process has been performed due to the text legibility aspect of the color design for the magazine cover.



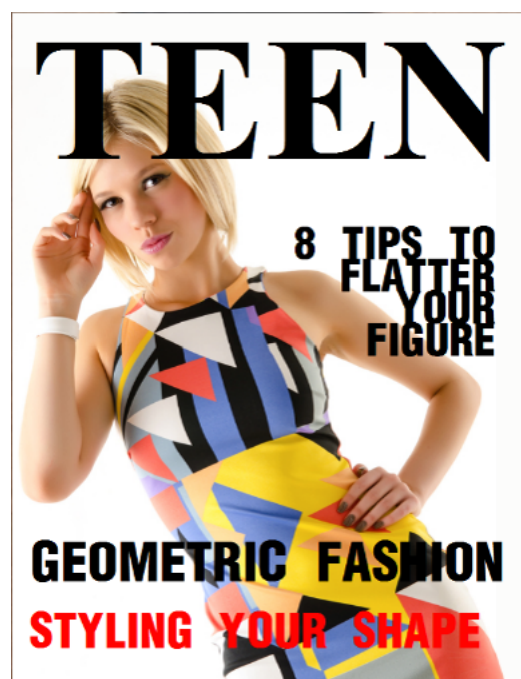
(a)



(b)



(c)



(d)

Fig. 5.14.: Automatic designs generated by ADoMC. Photos courtesy of Sean Molin (a, b, and c), and Martin O'Connell (d).

Note that the color modification process is part of our automatic color design solution, and there is no modification made by the user (refer to Sec. 5.7.3 for details). Also note that, as the results suggest, choosing the right content for a magazine cover is very important in the formation of an aesthetically compelling magazine cover. By content we mean the words that are chosen to be presented on the cover. The content of a magazine cover may be influenced by the hot topics and contemporary subjects in societies [260].

5.9 Conclusion and Future Work

At this stage of our work, the feedback of professional designers suggests that our automatic designs are aligned with their intuition. However, there are some limitations in our system which deserve more investigation for quantifying aesthetics of design. For instance, in our research we assume that the input to the system is a well-composed image. At this stage of the work, we consider a well-composed image to be an image consistent with the rule-of-thirds. To be considered well-composed, an image should also satisfy criteria related to other features, such as content and color combination. These features need to be addressed in our future work. Neither does our system at the moment take care of the different sizes of an input image. It assumes that the image is within the aspect ratio of the magazine cover to be designed. How to crop an image to obtain such an aspect ratio is another challenge under the umbrella of image resizing and image retargeting research. A further consideration is visual saliency. In some cases, Graph-based visual saliency (GBVS) does not satisfy our requirement for the salient part. We can instead use visual saliency to extract foreground from background. Another feature of design is making decisions about the choice of typeface. In current results, we have used a number of predefined typefaces. Although these typefaces were recommended to us by professionals for the purpose of design of a typical magazine cover, we are interested to determine what kind of typefaces can be used within the style of a magazine.

The style of a magazine, which is heavily influenced by the genre of magazine, has a major influence on the design of magazine covers. At this stage, our implementation does

not consist of design based on style or even genre of magazine. However, we have done some preliminary work about this aspect of design and part of our results are presented in the next chapter. In the next chapter, we present how color semantics can help us to recommend a design based on the mood that the user wants to convey in his/ her design.

6. RECOMMENDATION SYSTEM FOR AUTOMATIC DESIGN

6.1 Introduction

In the information era, the representation of information via visual media is a desire for both personal and business means, though consumers may prefer not to hire a professional to create designs for them. Personal communication through internet web services is becoming an essential part of contemporary life. New generations of communication devices also contribute to the need for the design of personal communication media. Hence, there is a great need and an increasing request for software applications that provide users with easy-to-follow steps for creating visual media, such as magazines, blogs, and other forms of self-publishing media. This need, in turn, inspires science and engineering to tackle the idea of visual media's automatic design.

An automatic design system, however, needs to take into account the key concepts of design and address the challenges that designers encounter when creating a design. The key concepts include: elements of design, principles in design, and aesthetics of design. By elements of design, we mean layout and visual balance considerations, typography, design of color, style, and context. Principles in design, however, refer to a number of rules that a designer may apply to accomplish the task of design. Aesthetics of design deals with the beauty of design and how good a design is, considering both its form and function. In fact, an automatic design system has to quantify these subjective concepts. Moreover, such a system has to perform the task of design automatically and autonomously. From a broader perspective, this system should be able to recommend to the user a number of designs and to customize the future designs for a user based on his/her personal preferences. Therefore, designing effective sets of interactions for user participation in the process of automatic design of visual media based on his/her preferences is required.

In this chapter, we introduce a recommendation system for automatic design of magazine covers. Professional designers start a design by developing a concept to convey the message or purpose of the design. Similarly, our system guides the user to understand the purpose of his/her design. The purpose of a magazine cover design relates to the context of its cover image [54]. The context of the cover image has several aspects, including the color mood and the objects. At this stage, we focus on the color mood: we recommend a set of semantic descriptors to the user to understand his/her preferred color mood in his/her design. The system then selects a number of photos from the user's album based on his/her preferences and generates alternative designs.

We focus on magazine covers as an application of the idea of automatic design, especially because of the need for self-publishing designs in recent years. This need has also been addressed in prior work, such as [258] and [112]. There are now several tools and services for the design of magazine covers. As an example, HP MagCloud websites [286] provide individuals as well as businesses with the ability to design and upload magazine covers to share with others. Moreover, there are several on-line tools and apps for magazine cover design for amateurs, which is in contrast with tools that facilitate professionals for the design of magazine covers. Consequently, some well-known magazines such as *Wired* and *Cosmopolitan* have started offering design of personalized magazine covers as promotions to customers. As an example of other applications of our work, consider the case for providing non-designer designers with the opportunity to design magazine-like documents to communicate with specific target audiences. For instance, Greig et al. [287] suggest a system which creates a travel brochure, but it has to be enhanced by adding features that recognize how to design an effective cover. Other potential users include small non-profit organizations that need to design magazine covers and other magazine-like documents to generate excitement and credibility about their work, without hiring professional designers [258]. Furthermore, as we suggest in the discussion section, our work could be considered to be a starting point to contribute to the idea of creativity support tools for both non-designer and designer users.

Our recommendation system is an extension of the automatic design system ADoMC, or Automatic Design of Magazine Covers (see the previous chapter), by Jahanian et al. [110]. Hence, we call our recommendation system R-ADoMC. Section 6.4 illustrates a schematic view of the system. We describe each part of R-ADoMC, including: Input User Interface in Sec. 6.5, which suggests different purposes of a design to the user; Evaluation of Input Photos in Sec. 6.6, which deals with the user's inputs; Designs User Interface in Sec. 6.7, which provides a rating mechanism of the generated designs for the user; and Personalization of Designs in Sec. 6.8, which recommends more designs based on the preferences of the user. We finally discuss the current limitations, challenges, and future work in Sec. 6.10. In the next section, we illustrate a common scenario typical for our users.

6.2 Scenario

Our users are non-designer designers, individuals and businesses who want to design without hiring a professional designer, while still wanting to create aesthetically compelling designs. Consider the following scenario: John has a pile of photos, and he wants to make a magazine cover with the best photo from the group. John also has a number of stories as his cover lines that he wants to place on the cover. He has thought about the priorities of his stories in terms of their importance. He knows what he would like to call his magazine title, but he may or may not have a predesigned masthead for his magazine title. John wants to make an attractive design, but he has no idea about the principles of design. The only thing that he knows is that he wants a "sporty and dynamic" design. Designers believe that if someone does not know how to design, he/she can easily make non-pleasing designs [260]. They also believe that any changes made in a design would lead to a different design, and the designer would have to start a new design from scratch. From the above input, R-ADoMC is implemented to create a design from start to finish. It does not allow for any modifications on the user's part. Instead, it provides a number of designs to the user and lets him/her choose one. It also recommends a design based on the preference of the user. In John's case, R-ADoMC takes his photos and stories (i.e. cover lines), and then it suggests

a number of adjectives to him that describe some possible kinds of design, e.g., “clean and clear”, “dynamic and active”, “elegant”, “formal”, etc. In the next sections, we discuss these adjectives, their origins, and the way that we deploy them to create a design.

6.3 Related Work

We believe the work of visual media’s automatic design, magazine covers in particular, is a multidisciplinary task that spans several different topics, including automated layout, color design, typography, computational aesthetics, and recommendation with personalization. In the previous chapter, we have reviewed the related work in each area and discussed how we apply or extend the prior work. The following reviews the prior work in design recommendation systems.

Recommending a generated design from scratch for the user is another contribution of this thesis. [114] discusses how designers can benefit from examples recommended by a recommending system to support their creativity in new designs. The main difference, however, is that in our work we create designs and recommend them to non-designer users. In recommendation systems, typically, prediction of a user’s preferences is based on a rating data set, accumulated from the same user and similar users [288]. In our work, we personalize the future designs for a user by predicting his/her personal preferences in design and by inferring a style for him/her based on his/her ratings.

6.4 Software Framework

This section describes how we think about the problem of automatic design from the viewpoint of software engineering and computer science. We endeavor to quantify the elements and concepts in design and bring these subjective topics to the realm of objection. Our approach includes close collaboration with professional designers, magazine art directors and editorial boards, and journalists. The current software framework is the projection of a number of rules, mathematical expressions, and algorithms delineated and devised from the lessons we have learned so far. The following sub-sections describe each part of

the system in more detail. We first start by looking at a schematic view of the framework and the relation of each part with others.

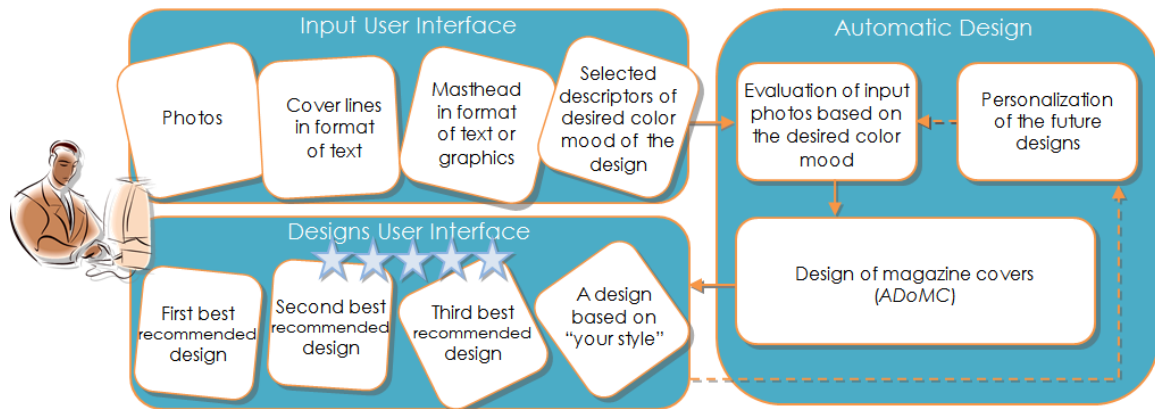


Fig. 6.1.: A schematic view of R-ADoMC. The user uploads his/her photos, cover lines, and masthead. Then, he/she chooses some recommended descriptors to indicate the desired color mood of his/her design. R-ADoMC evaluates the user's photos and chooses the best. It then sends the selected photos to the design engine. The generated designs will be shown to the user, who can rate the generated designs. The ratings will be considered for the user's future designs to recommend him/her one more design based on his/her style.

6.4.1 Schematic View

As Fig. 6.1 illustrates, the system is comprised of three major parts: Input User Interface, Automatic Design Engine, and Designs User Interface. The system first takes the user's inputs, which includes a collection of his/her photos, a number of text strings as his/her magazine cover stories (e.g. headline and byline), a masthead for the title in the format of a text string or a predefined logo, and a number of predefined adjectives describing the desired color mood of his/her design. The system then evaluates the user's photos based on his/her chosen adjectives. In this way, the system can order the input photos based on their aesthetics, including both form and functional aspects. The best photos are then sent to the design engine. This engine is called ADoMC and is based on work by Jahanian et

al. [110]. The designs generated by ADoMC, which in turn pass a visual composition criterion, are presented as the final designs to the user by listing them in order of best design. The user can rate the generated designs using a five-star rating scale. The system incorporates the user's ratings in the personalization part to customize his/her future designs. In this fashion, the system will suggest to the user one more design among the other designs that he/she may like by indicating that it is based on "your style."

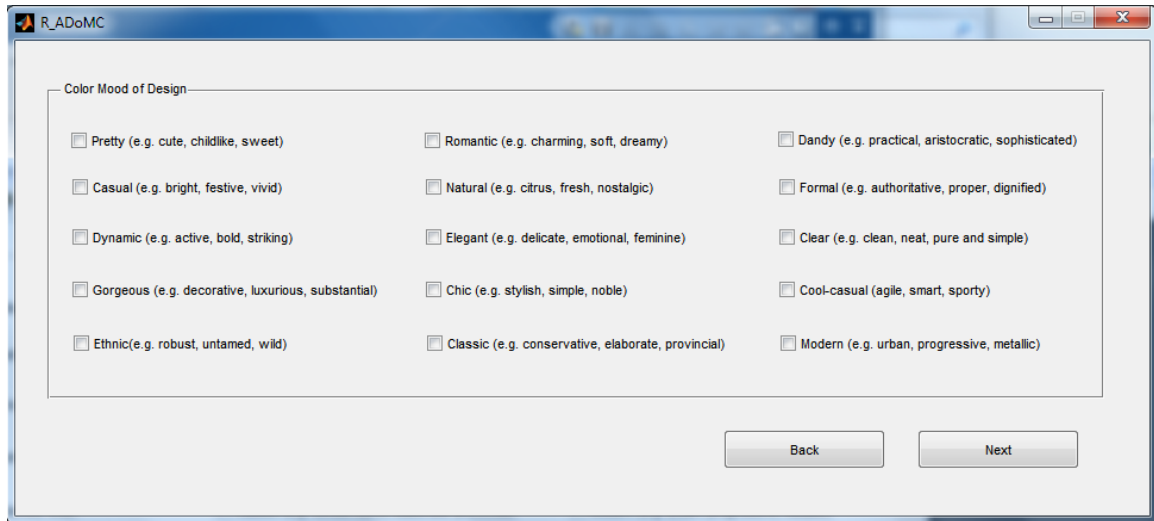


Fig. 6.2.: Descriptors for color mood of design. The user can tell R-ADoMC which color mood he/she wishes to convey in his/her design. These descriptors are suggested by [43] to relate the world of colors to the world of styles and human preferences.

6.5 Input User Interface

A magazine cover has a main photo called a cover image, a title called a masthead, and a number of stories, including headline, byline, and other cover lines [54]. The user can upload his/her cover image, masthead, and cover lines to R-ADoMC. Users generally have a number of photos and may wonder which photo could be a good candidate for a magazine cover. On the other hand, since the context of the cover image influences the purpose of a magazine cover design, we recommend a list of adjectives to the user to describe several

color moods in the context of a magazine cover. Although the context of a photo includes aspects other than color, at this stage of our work, we address only the color aspect. After the user chooses a set of adjectives, or color mood descriptors, R-ADoMC can evaluate the user's photos, rate them, and generate alternative designs. Figure 6.2 illustrates the menu of the color mood descriptors in R-ADoMC.

6.6 Evaluation of Input Photos

Computational aesthetics of photos is one of the recent interests in computer science and engineering. However, most of the related work is focused on a number of low level features in a photo, such as color distribution, texture, and objects [198]. These features are part of the aesthetics of photos, but they are context-free. Since a design has a mission and conveys a purpose to the audience, we also need a set of high level features to be able to describe a design and its color mood. Our solution to this challenge is inspired by the work on the Color Image Scale by Kobayashi [43] (see Fig. 6.3). Kobayashi has proposed a semantic view to the aesthetics of colors to relate the world of colors to the world of objects, life styles, and human preferences. Therefore, we use his work to quantify the mood of our designs from a semantic viewpoint.

Kobayashi's Color Image Scale is an attempt to investigate the meanings that people may perceive from the colors and color combinations, from a number of psychological, systematic, and pragmatic approaches. This viewpoint comes from the idea of associating the physical attributes of the colors – hue, value, and chroma — to a higher level of abstract attributes of colors – warm-cool and soft-hard. Kobayashi and his institute, Nippon Color and Design Research, have conducted more than a decade of psychophysical experiments on colors. They asked individuals and groups of people to match a color and a combination of three colors with an adjective. Through factor analysis, a semantic differential method, and cluster analysis, Kobayashi came up with a list of adjectives as color descriptors or “color-images.” Kobayashi then applied his scale to different color designs, for example, clothes, food, and several other industrial design applications [42]. Kobayashi

later published his Color Image Scale in a book with the same name [43]. Today a Japanese website uses his work in a search engine for images [289]. Recently [150] has deployed Kobayashi's scale for image retrieval.

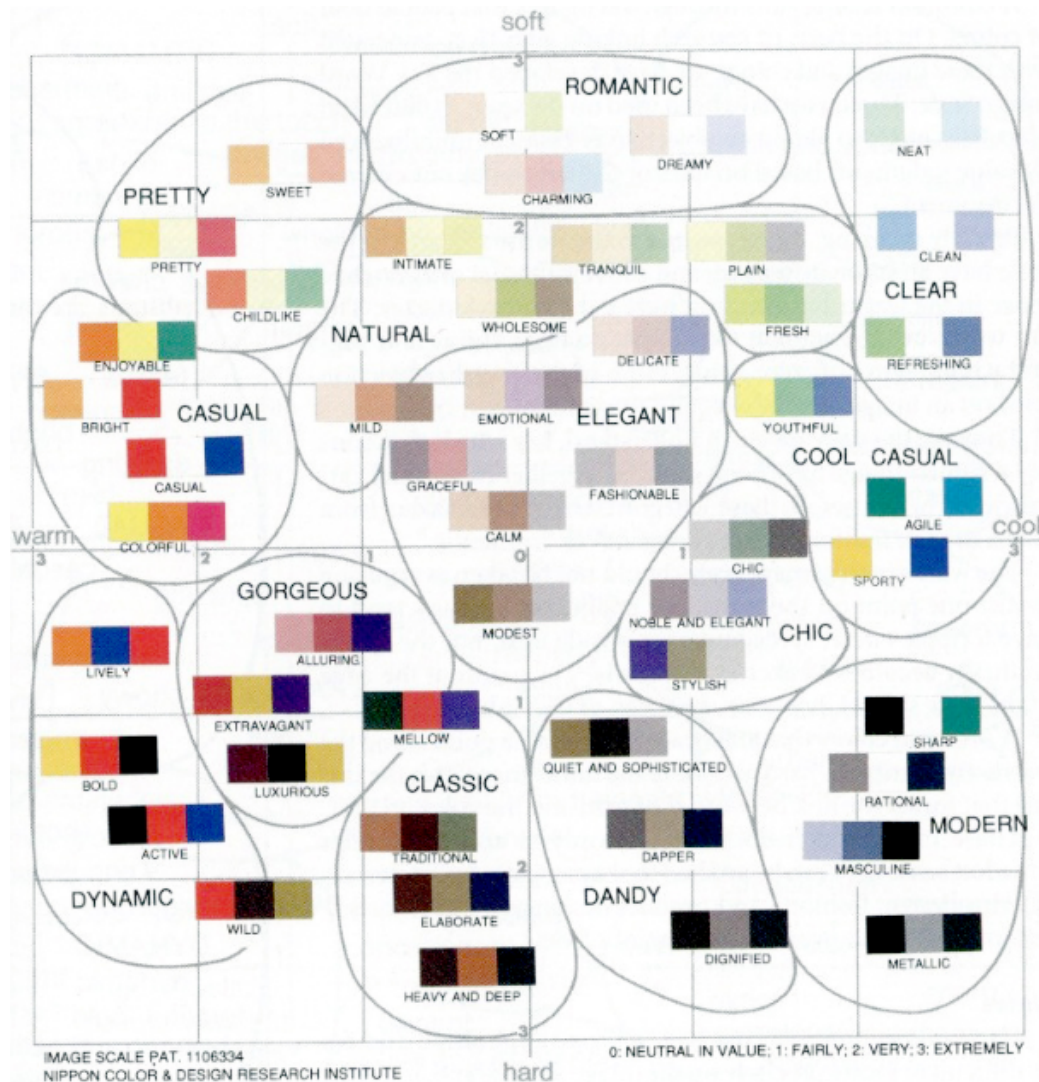


Fig. 6.3.: Kobayashi's Color Image Scale. Image taken from [43]. The contours show 15 patterns, labeled by all-capital letters. Note that Kobayashi has defined 1170 3-color combinations on this scale, however for simplicity only some of them are shown. We use these patterns to describe the desired color mood of the design.

We deploy the mathematical framework that [150] develops to implement Kobayashi's color scale for the purpose of image retrieval; hence, we use almost the same notation. We, however, use this framework to measure and evaluate the input photos for the purpose of a design. Kobayashi first defines 130 Basic Colors. He then defines 1170 3-Color Combinations and associates them with 180 adjectives (called color-images) or Labels. Each 3-Color Combination is also associated with a five-star rating that indicates the frequency of its usage. The Labels are later classified into 15 Patterns. The Patterns represent a number of selected terms in fashion and lifestyle. Kobayashi positions his 3-Color Combinations on a scale of warm-cool and soft-hard, and based on this arrangement, defines the boundaries of each Pattern. Figure 6.3 illustrates Kobayashi's Color Image Scale with the 15 Patterns and some of his 3-Color Combinations.

Our problem is how to locate a given image on Kobayashi's Color Image Scale. First we need to find a way to represent an image in sRGB color space in terms of Kobayashi's 130 Basic Colors. One way is to represent the image by its histogram. In other words, the problem is: how to convert the sRGB histogram of an image to its corresponding histogram in Kobayashi's 130 Basic Colors. To do so, we quantize the sRGB color space to a smaller space with 8 levels of R, 8 levels of G, and 8 levels of B. This leads to 512 sRGB colors, or bins. We obtain the Kobayashi's histogram of an image, denoted by h_K from the sRGB histogram of the image, denoted by h_{RGB} , according to the following equation:

$$h_K = h_{RGB} \cdot T \quad (6.1)$$

Here T is the transformation matrix. We construct T by first converting both the quantized sRGB space and the 130 Basic Colors to CIELab color space. This space is perceptually uniform, and so we measure the distance between each pair of colors converted from these two color spaces in terms of their Euclidean distance in CIELab, denoted by ΔE_r . We can then define the elements of the transformation matrix T :

$$t_{rl} = \begin{cases} 1 - \frac{\Delta E_r}{10} + 0.1 & \text{if } \Delta E_r \leq 10 \\ 0 & \text{otherwise} \end{cases} \quad (6.2)$$

where $r = 1, \dots, 512$, and $l = 1, \dots, 130$.

Hence, we obtain Kobayashi's histogram h_K of an image. The next step is computation of the distribution of any 3-Color Combination in the given image from its corresponding h_K . To do so, we first define the following notation to refer to a 3-Color Combination:

$$c_n^{(3)} = \{b_n^1, b_n^2, b_n^3\} \quad (6.3)$$

where $0 \leq b_n^i \leq 130$ denotes the index of a Basic Color.

Now we define h_σ as the distribution of the all 3-Color Combinations in a given image by:

$$h_\sigma = [\sigma_1, \dots, \sigma_{1170}]^T \quad (6.4)$$

where

$$\sigma = \min(h_K(b_n^1), h_K(b_n^2), h_K(b_n^3)) \quad (6.5)$$

for $n = 1, \dots, 1170$.

Next, we define matrix L as an index matrix to associate the 180 Labels to the 1170 3-Color Combinations:

$$L = [\mathbf{l}_1 \dots \mathbf{l}_{180}] \quad (6.6)$$

where L is a 1170×180 matrix and

$$l_{ij} = \begin{cases} 1 & \text{if label } j \text{ is associated to Color Combination } i \\ 0 & \text{otherwise} \end{cases} \quad (6.7)$$

Also, we define the diagonal matrix R , which contains the ratings of each 3-Color Combination. Hence, we obtain the distribution of each Label in an image by h_L as:

$$h_L = h_\sigma \cdot R \cdot L \quad (6.8)$$

Similar to Eqs 6.6 and 6.7, we can define the matrix U as an index matrix to associate the 15 Patterns to the 180 Labels. Hence, we obtain the distribution of each Pattern in an image as follows:

$$P = h_L \cdot U \quad (6.9)$$

Finally, given the location of all the 180 Labels on the scale from the matrix E , we can obtain s , the location of the given image on the scale from:

$$s = h_L \cdot E^T \quad (6.10)$$

As an example, Fig. 6.4 illustrates the location of an image on the Color Image Scale. In fact, we can find which of the Patterns, or in our word, color moods an image holds. The normalized values of matrix P give us a measure of closeness or similarity of an image to each Pattern. We use this measure to rate the user's photos and compute the best matches to Patterns. As we mentioned, this method deals with the high level features of an image's aesthetics. However, we also consider a low level feature in images, which is the layout composition of the image. We suggest this feature because the composition of the cover image heavily influences the visual balance and layout of the design of a magazine cover. Therefore, if an image maintains a busy composition, it will not be included among the final designs. At this stage, we equate the layout composition of an image with its visual saliency and compute it with the Graph-Based Visual Saliency algorithm [164].

6.7 Designs User Interface

Getting the user to effectively visualize the results is a challenge when designing a user interface for recommendation systems. This, in turn, influences the personalization of the user's results. In our problem, we aim to present more than one design to the user in order to understand his/her design's purpose and to customize his/her future designs. We believe this is an added value for non-designers because it provides them designs that are aligned with their personal preferences and styles. In this fashion, we devise an interactive process

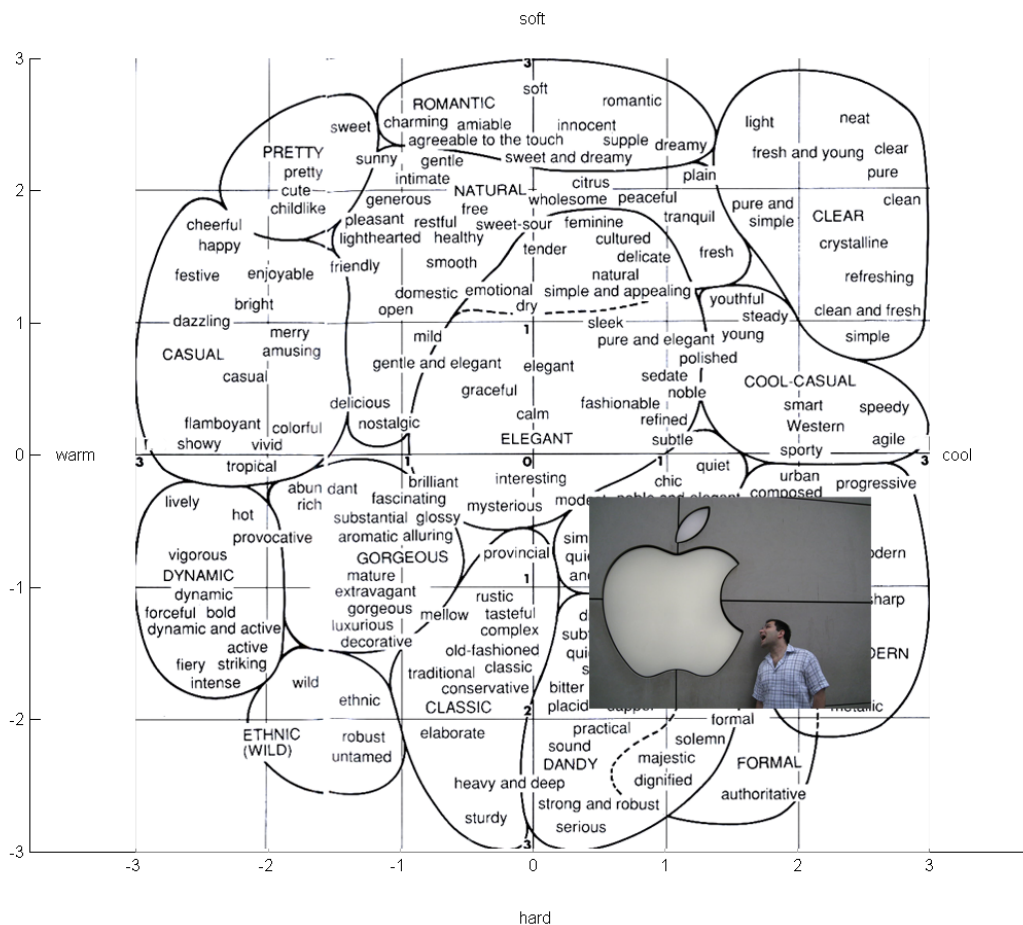


Fig. 6.4.: The location of an image on Kobayashi's Color Image Scale. This image is considered to be CHIC and MODERN according to this scale. The scale is taken from [43].

between the automatic system and the user while creating a design, which also happens in practice between a professional designer and his/her client. Figure 6.5 illustrates R-ADoMC's user interface. R-ADoMC displays the first choice design in the main GUI and arranges two more design choices, along with a design based on the user's accumulated ratings in a separate panel. It also provides a five-star rating scale option by the user for these designs [290] provides a number of guidelines for choosing a rating scale, the options being unary ("like it"), binary (thumbs up/thumbs down), five-star, or a 100-point slider. The degree of accuracy and user friendliness of the five-star method is a good choice for

our work. In this way, we accumulate a better understanding of what kind of design the user likes, and consequently, we can customize his/her future designs.

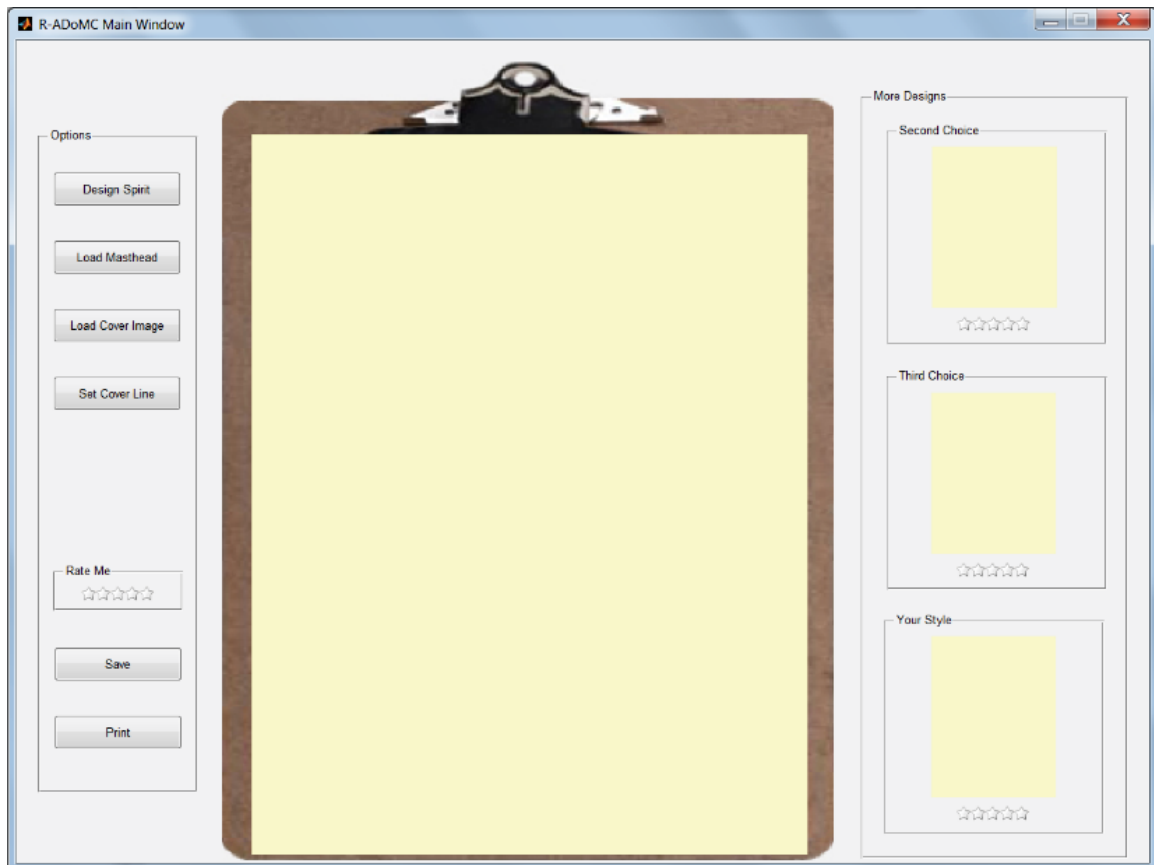


Fig. 6.5.: Designs UI for visualization of the designs generated in R-ADoMC. A rating option is available for each design, and the ratings will be considered in the user's future designs. The "your style" design is a representation of the idea of personalized designs for the user. The more that the user works with the system, the more that the system is able to personalize the future designs.

6.8 Personalization of Designs

Is it legitimate to personalize designs? Although form follows function – or the context of a design should be aligned with the mission of the design – it can also be personalized. In fact, one way of thinking about personalization of a design is to think about the design's

style, which comes from the preferences of a designer. There is a consensus between artists that each professional has his/her own style or signature, making some of the designers identifiable based on their designs even without their explicit identification. On the other hand, there might be alternative good designs for a specific purpose, and choosing one from several, ties with the personal preferences of the user. Hence, providing alternative designs and allowing a user to choose one design based on his/her personal preferences is an added value to the function of automatic design systems. Accordingly, R-ADoMC implements the idea of design personalization to support the user. As the user works more with R-ADoMC and rates the generated designs, the rating data is accumulated for future designs.

There are three different approaches for recommending an item in current recommendation systems, content-based, collaborative filtering, and a combination of both these known as hybrid [288]. At this stage, our system does not consider any collaborative filtering because its personalization procedure is single-user based rather than multiple-user based. As Fig. 6.5 illustrates, the system suggests the personalized design as an extra design, labeled as “your style.” To anticipate this personalized design from the previous ratings of the user, we define the prediction of the ratings of the generated designs as a linear regression problem. To solve this problem, we use the LASSO algorithm [291] with L_1 regularization.

6.9 Experimental Results

Figure 6.6 illustrates several ordered results. Generally speaking, the results generated by ADoMC are acceptable to our professional designers. However, a set of psychophysical experiments to evaluate the results by users is part of our future work. The system is implemented in Matlab and takes a few minutes to present the results (for a photo album of about 100 images), on a 64-bit PC machine with a 2-core 3GHz Intel processor and 4GB of RAM.

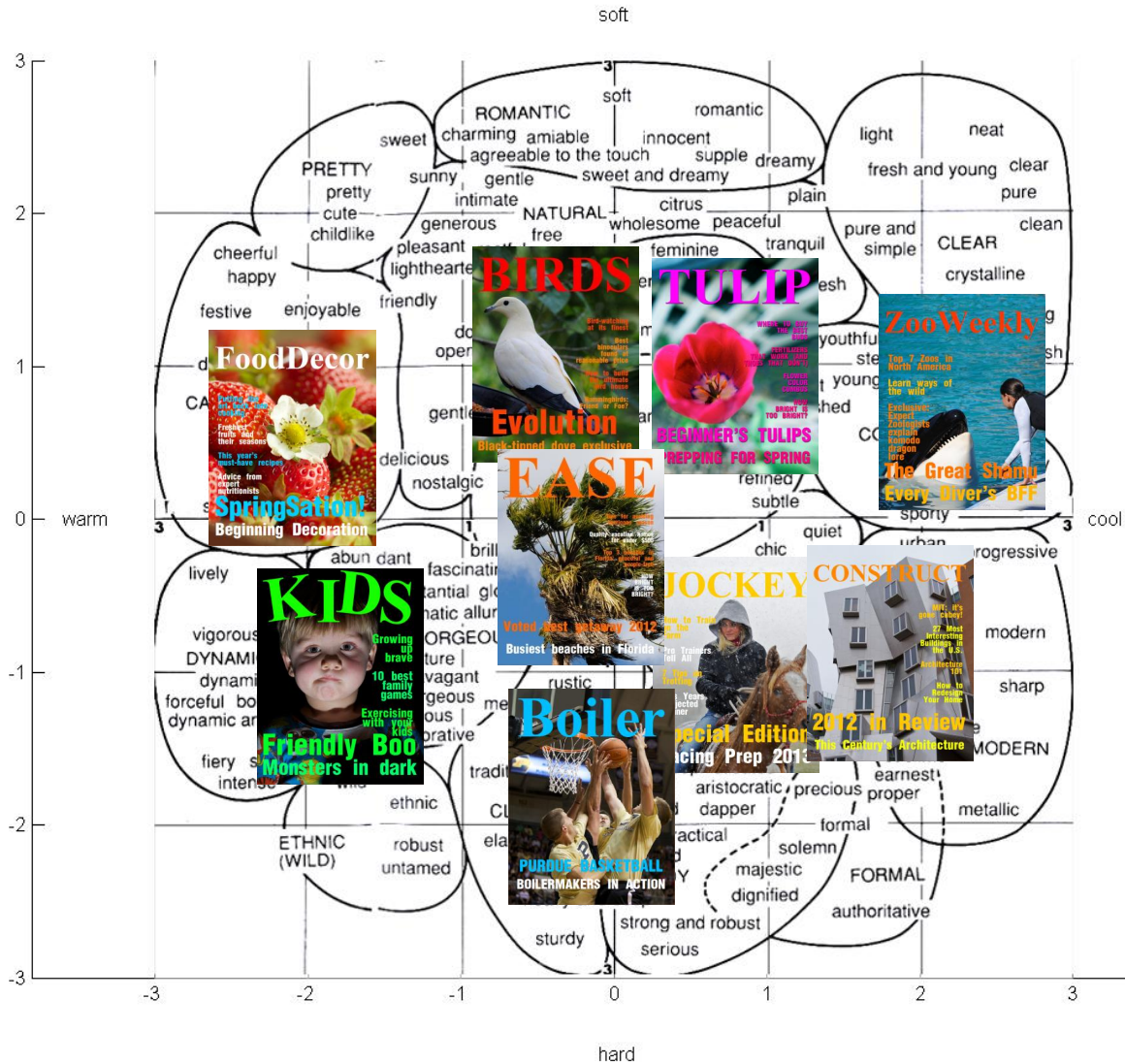


Fig. 6.6.: Illustrations of magazine covers that were automatically designed on Kobayashi's Color Image Scale (the scale image is taken from [43]). The cover images of these designs have been selected according to Kobayashi's Patterns to represent the mood of the magazine cover designs. Images reproduced from [138]. Photos courtesy of Andy Jessop.

6.10 Conclusion and Future Work

R-ADoMC is a recommendation system for the automatic design of magazine covers, and the purpose of this system is to aid the non-designer designer in creating aesthetically

compelling designs. After a user uploads his/her photos, the system first suggests a number of semantic features, e.g., “elegant,” “dynamic and active,” or “sporty,” to describe the color mood for the purpose of his/her design. Based on these high level features and a number of low level features, such as the complexity of a photo’s visual balance, our system then selects the most suitable photos from the user’s album for his/her design. Following this, R-ADoMC generates several alternative designs that can be rated by the user. Consequently, the system generates future designs grounded in the user’s style, thus personalizing the designs based on his/her preferences.

The feedback that we have received from professional designers suggests that our automatic designs are aligned with their intuition. However, there is a need to devise a solid framework to evaluate R-ADoMC’s generated results. This problem is highly correlated with aesthetics measurements. Quantifying aesthetics is still a challenge in computer science and engineering. In our case, while seeking a good solution by discussing this problem with our partners in the School of Visual and Performing Arts at Purdue University, we found that there might be several alternative good designs. So, we might be able to restrict the problem to what is a bad design. That inspired us to follow the principles of design. We believe that our work, among other previous work, emphasizes the need of devising and designing a set of measures and user experiments to evaluate aesthetics of the visual design. It is also worthwhile to mention that one dimension of aesthetics is beauty, which might be perceived through personal interest and experience, as suggested by Frohlich in “Beauty as a Design Prize” [292] and also by Barthes [293]. This viewpoint supports the notion of personalized design, or “Your Style,” which, as suggested in this chapter, may leverage the transcendental experience of beauty. Moreover, as Frohlich suggests, the perception of beauty can be stimulated by the inclusion of a more subjective experience in design. This is another support for the idea of recommendation for automatic design, as it considers the user’s preferences. In fact, in interaction design, engaging the non-designer designer in the process of a design’s creation by having a personal experience can arouse the experience of beauty, and hence, aesthetics.

Although our framework provides personalized designs, we endeavor to develop a framework which generates designs that are effective in both the function and form of aesthetics. Keeping this in mind, there are a number of limitations in R-ADoMC that must be addressed. One limitation is that we reduce the dimensionality of a design's purpose to the color mood dimension, as well as visual complexity of the cover image. Although these two dimensions are important and influence the aesthetics of a design heavily, there exist other dimensions such as the objects in the cover image. This correlates with the context evaluations of the photos, and so these contextual factors need to be further investigated. Another limitation includes the hypothetical case that a user has an album of similar photos, e.g., all the photos are classified as modern in his/her album. While our system would provide several designs, we still require another feature or set of features to identify the ranking of these designs based on what the user needs. An additional limitation is that we do not consider specific typefaces for our typography, and this problem ties with the class of magazine, which contributes to defining a magazine's style. We aim to address design based on the class of magazines within our future work.

Our current system, R-ADoMC, is a software framework meant to emulate the process that a professional designer attends to when creating a design. In fact, one of the key components of this framework is developing an interactive process between the software and the user to understand what kind of design the user needs or wants. As also indicated by Swearingen et al. [294], understanding the user's needs is important in designing the interactions of recommendation systems. On the other hand, designing interactions for recommendation systems can also be expanded to contribute to the idea of creativity support tools. Creativity support tools facilitate users to "explore, discover, imagine, innovate, compose, and collaborate" [295]. However, as indicated by Shneiderman [188], creative work often begins by viewing previous projects and similar samples. Since R-ADoMC recommends alternative designs to the user based on his/her preferences, it can be an underlying framework for a creativity support system in visual design. If the user is unaware of key concepts in design or principles of design, it is an added value to recommend these semantic descriptors that make sense to him/her and help to describe a design. It is, how-

ever, more ambitious to escalate the creativity of our non-designer user in the design of visual media. This desire requires a series of effective interactions between the user and the system to include the user in the creative process of his/her design. Hence, we endeavor to evolve R-ADoMC's framework into a creativity support tool for the process of visual design and concept design for media.

REFERENCES

REFERENCES

- [1] S. Bødker, “When second wave hci meets third wave challenges,” in *Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles*. ACM, 2006, pp. 1–8.
- [2] L. E. Udsen and A. H. Jørgensen, “The aesthetic turn: unravelling recent aesthetic approaches to human-computer interaction,” *Digital Creativity*, vol. 16, no. 04, pp. 205–216, 2005.
- [3] J. K. McDonald, “The creative spirit of design,” *TechTrends*, vol. 55, no. 5, pp. 53–58, 2011.
- [4] C. Korsmeyer, “Aesthetics: The big questions,” 1998.
- [5] D. A. Norman, *Emotional design: Why we love (or hate) everyday things*. Basic books, 2004.
- [6] J. Frascara, *Communication design: principles, methods, and practice*. Allworth Communications, Inc., 2004.
- [7] Q. Newark, *What is graphic design?* Rockport Publishers, 2007.
- [8] M. Barnard, *Graphic design as communication*. Routledge, 2013.
- [9] H. Zettl, *Sight Sound and Motion*. Cengage Learning, 2012.
- [10] R. E. Horn, *Visual language*. MacroVU, Incorporated, 1998.
- [11] T.-M. Karjalainen, *Semantic transformation in design: Communicating strategic brand identity through product design references*. University of Art and Design Helsinki, 2004.
- [12] M. Barnard, *Graphic design as communication*. Psychology Press, 2005.
- [13] R. Hoffmann and K. Krauss, “A critical evaluation of literature on visual aesthetics for the web,” in *Proceedings of the 2004 annual research conference of the South African institute of computer scientists and information technologists on IT research in developing countries*. South African Institute for Computer Scientists and Information Technologists, 2004, pp. 205–209.
- [14] F. Klett, “Visual communication in web-based learning environments,” *Educational Technology & Society*, vol. 5, no. 4, pp. 38–48, 2002.
- [15] D. A. Norman, *Emotional design: Why we love (or hate) everyday things*. Basic books, 2007.

- [16] G. Lindgaard, G. Fernandes, C. Dudek, and J. Brown, "Attention web designers: You have 50 milliseconds to make a good first impression!" *Behaviour & information technology*, vol. 25, no. 2, pp. 115–126, 2006.
- [17] G. Lindgaard, C. Dudek, D. Sen, L. Sumegi, and P. Noonan, "An exploration of relations between visual appeal, trustworthiness and perceived usability of homepages," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 18, no. 1, p. 1, 2011.
- [18] N. Tractinsky, A. Cokhavi, M. Kirschenbaum, and T. Sharfi, "Evaluating the consistency of immediate aesthetic perceptions of web pages," *International Journal of Human-Computer Studies*, vol. 64, no. 11, pp. 1071–1083, 2006.
- [19] J. Itten, *The Art of Color*. Wiley, 1974.
- [20] G. D. Birkhoff, *Aesthetics Measure*. Harvard University Press, 1933.
- [21] D. C. L. Ngo, L. S. Teo, and J. G. Byrne, "A mathematical theory of interface aesthetics," in *Visual mathematics*, vol. 2, no. 4. Mathematical Institute SASA, 2000.
- [22] D. C. L. Ngo, A. Samsudin, and R. Abdullah, "Aesthetic measures for assessing graphic screens," *J. Inf. Sci. Eng.*, vol. 16, no. 1, pp. 97–116, 2000.
- [23] J. Rigau, M. Feixas, and M. Sbert, "Informational aesthetics measures," *IEEE Computer Graphics and Applications*, vol. 28, no. 2, pp. 24–34, 2008.
- [24] P. Machado and A. Cardoso, "Computing aesthetics," in *Advances in Artificial Intelligence*. Springer, 1998, pp. 219–228.
- [25] G. R. Greenfield, "Simulated aesthetics and evolving artworks: A coevolutionary approach," *Leonardo*, vol. 35, no. 3, pp. 283–289, 2002.
- [26] T. Lavie and N. Tractinsky, "Assessing dimensions of perceived visual aesthetics of web sites," *International journal of human-computer studies*, vol. 60, no. 3, pp. 269–298, 2004.
- [27] N. Damera-Venkata, J. Bento, and E. O'Brien-Strain, "Probabilistic document model for automated document composition," in *Proceedings of the 11th ACM symposium on Document engineering*. ACM, 2011, pp. 3–12.
- [28] R. Datta, D. Joshi, J. Li, and J. Wang, "Studying aesthetics in photographic images using a computational approach," *Computer Vision–ECCV 2006*, pp. 288–301, 2006.
- [29] M. Bense, *Einführung in die informationstheoretische Ästhetik: Grundlegung und Anwendung in der Texttheorie*. Rowohlt [Taschenbuch-Verlag, 1969, vol. 320.
- [30] G. Greenfield, "On the origins of the term" computational aesthetics"." in *Computational Aesthetics*, 2005, pp. 9–12.
- [31] J. Rigau, M. Feixas, M. Sbert, and C. Wallraven, "Toward Auvers period: evolution of van gogh's style," in *Proceedings of the Sixth international conference on Computational Aesthetics in Graphics, Visualization and Imaging*. Eurographics Association, 2010, pp. 99–106.
- [32] D. E. Berlyne, *Studies in the new experimental aesthetics: Steps toward an objective psychology of aesthetic appreciation*. Hemisphere, 1974.

- [33] S. J. Harrington, J. F. Naveda, R. P. Jones, P. Roetling, and N. Thakkar, "Aesthetic measures for automated document layout," in *Proceedings of the 2004 ACM symposium on Document engineering*. ACM, 2004, pp. 109–111.
- [34] S. Lok, S. Feiner, and G. Ngai, "Evaluation of visual balance for automated layout," in *Proceedings of the 9th international conference on Intelligent user interfaces*. ACM, 2004, pp. 101–108.
- [35] H. Balinsky, "Evaluating interface aesthetics: measure of symmetry," in *Electronic Imaging 2006*. International Society for Optics and Photonics, 2006, pp. 607 608–607 608.
- [36] A. Klinger and N. A. Salingaros, "A pattern measure," *arXiv preprint arXiv:1108.5508*, 2011.
- [37] G. R. Greenfield, "Evolving aesthetic images using multiobjective optimization," in *Evolutionary Computation, 2003. CEC'03. The 2003 Congress on*, vol. 3. IEEE, 2003, pp. 1903–1909.
- [38] B. J. Ross, W. Ralph, and H. Zong, "Evolutionary image synthesis using a model of aesthetics," in *Evolutionary Computation, 2006. CEC 2006. IEEE Congress on*. IEEE, 2006, pp. 1087–1094.
- [39] J. Itten, *The elements of color*. John Wiley & Sons, 1970, vol. 4.
- [40] M. Tokumaru, N. Muranaka, and S. Imanishi, "Color design support system considering color harmony," in *Fuzzy Systems, 2002. FUZZ-IEEE'02. Proceedings of the 2002 IEEE International Conference on*, vol. 1. IEEE, 2002, pp. 378–383.
- [41] D. Cohen-Or, O. Sorkine, R. Gal, T. Leyvand, and Y. Xu, "Color harmonization," in *ACM Transactions on Graphics (Proc. of ACM SIGGRAPH)*, vol. 25, no. 3. ACM, 2006, pp. 624–630.
- [42] S. Kobayashi, "The aim and method of the color image scale," *Color Research & Application*, vol. 6, no. 2, pp. 93–107, 1981.
- [43] S. Kobayashi and L. Matsunaga, *Color image scale*. Kodansha international Tokyo, 1991.
- [44] R. Arnheim, *Art and visual perception: A psychology of the creative eye*. Univ. of California Press, 1954.
- [45] P. J. Locher, "An empirical investigation of the visual rightness theory of picture perception," *Acta Psychologica*, vol. 114, no. 2, pp. 147–164, 2003.
- [46] C. W. Tyler, "Some principles of spatial organization in art," *Spatial Vision*, vol. 20, no. 6, pp. 509–530, 2007.
- [47] I. McManus, K. Stöver, and D. Kim, "Arnheim's gestalt theory of visual balance: Examining the compositional structure of art photographs and abstract images," *i-Perception*, vol. 2, no. 6, p. 615, 2011.
- [48] M. Moshagen and M. T. Thielsch, "Facets of visual aesthetics," *International Journal of Human-Computer Studies*, vol. 68, no. 10, pp. 689–709, 2010.

- [49] L. Eisemann, *Pantones Guide to Communicating with Color*. Grafix Press, Ltd, 2000.
- [50] M. M. Aslam, “Are you selling the right colour? a cross-cultural review of colour as a marketing cue,” *Journal of marketing communications*, vol. 12, no. 1, pp. 15–30, 2006.
- [51] G. J. Gorn, A. Chattopadhyay, T. Yi, and D. W. Dahl, “Effects of color as an executional cue in advertising: they’re in the shade,” *Management Science*, vol. 43, no. 10, pp. 1387–1400, 1997.
- [52] L.-C. Ou, M. Ronnier Luo, P.-L. Sun, N.-C. Hu, H.-S. Chen, S.-S. Guan, A. Woodcock, J. L. Caivano, R. Huertas, A. Treméau *et al.*, “A cross-cultural comparison of colour emotion for two-colour combinations,” *Color Research & Application*, vol. 37, no. 1, pp. 23–43, 2012.
- [53] A. Whitfield and J. Wiltshire, “Design training and aesthetic evaluation: An intergroup comparison,” *Journal of Environmental Psychology*, vol. 2, no. 2, pp. 109–117, 1982.
- [54] C. Foges, *Magazine Design*. RotoVision Press, Switzerland, 1999.
- [55] D. M. Blei, A. Y. Ng, and M. I. Jordan, “Latent Dirichlet allocation,” *the Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003.
- [56] L. Shu, B. Long, and W. Meng, “A latent topic model for complete entity resolution,” in *Data Engineering, 2009. ICDE’09. IEEE 25th International Conference on*. IEEE, 2009, pp. 880–891.
- [57] S. Lin and P. Hanrahan, “Modeling how people extract color themes from images,” in *ACM Human Factors in Computing Systems (CHI)*, 2013.
- [58] Z. Pylyshyn *et al.*, “Is vision continuous with cognition? the case for cognitive impenetrability of visual perception,” *Behavioral and brain sciences*, vol. 22, no. 3, pp. 341–365, 1999.
- [59] P. Skorupski and L. Chittka, “Is colour cognitive?” *Optics & Laser Technology*, vol. 43, no. 2, pp. 251–260, 2011.
- [60] G. Derefeldt, T. Swartling, U. Berggrund, and P. Bodrogi, “Cognitive color,” *Color Research & Application*, vol. 29, no. 1, pp. 7–19, 2004.
- [61] G. W. Humphreys and V. Bruce, *Visual cognition: Computational, experimental, and neuropsychological perspectives*. Lawrence Erlbaum Associates, Inc, 1989.
- [62] L. W. Barsalou, “Perceptual symbol systems,” *Behavioral and brain sciences*, vol. 22, no. 04, pp. 577–660, 1999.
- [63] J. Heer and M. Stone, “Color naming models for color selection, image editing and palette design,” in *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*. ACM, 2012, pp. 1007–1016.
- [64] B. Berlin, *Basic color terms: Their universality and evolution*. University of California Pr, 1991.

- [65] P. Kay and C. K. McDaniel, "The linguistic significance of the meanings of basic color terms," *Language*, pp. 610–646, 1978.
- [66] S. E. Palmer, *Vision science: Photons to phenomenology*. The MIT press, 1999.
- [67] C. E. Osgood, "The nature and measurement of meaning," *Psychological bulletin*, vol. 49, no. 3, p. 197, 1952.
- [68] C. E. Osgood, *The measurement of meaning*. University of Illinois Press, 1957, vol. 47.
- [69] C. E. Osgood, "The cross-cultural generality of visual-verbal synesthetic tendencies," *Behavioral Science*, vol. 5, no. 2, pp. 146–169, 1960.
- [70] C. E. Osgood, "Exploration in semantic space: A personal diary1," *Journal of Social Issues*, vol. 27, no. 4, pp. 5–64, 1971.
- [71] F. M. Adams and C. E. Osgood, "A cross-cultural study of the affective meanings of color," *Journal of Cross-Cultural Psychology*, vol. 4, no. 2, pp. 135–156, 1973.
- [72] B. Wright and L. Rainwater, "The meanings of color," *The Journal of General Psychology*, vol. 67, no. 1, pp. 89–99, 1962.
- [73] P. Green-Armytage, "The value of knowledge for colour design," *Color Research & Application*, vol. 31, no. 4, pp. 253–269, 2006.
- [74] L.-C. Ou, M. R. Luo, A. Woodcock, and A. Wright, "A study of colour emotion and colour preference. part I: Colour emotions for single colours," *Color Research & Application*, vol. 29, no. 3, pp. 232–240, 2004.
- [75] L.-C. Ou, M. R. Luo, A. Woodcock, and A. Wright, "A study of colour emotion and colour preference. part II: Colour emotions for two-colour combinations," *Color Research & Application*, vol. 29, no. 4, pp. 292–298, 2004.
- [76] C. L. Hardin and L. Maffi, *Color categories in thought and language*. Cambridge University Press, 1997.
- [77] M. Thüring and S. Mahlke, "Usability, aesthetics and emotions in human–technology interaction," *International Journal of Psychology*, vol. 42, no. 4, pp. 253–264, 2007.
- [78] S. Lin, J. Fortuna, C. Kulkarni, M. Stone, and J. Heer, "Selecting semantically-resonant colors for data visualization," in *Eurographics Conference on Visualization (EuroVis), 2013 15th Annual Visualization Conference on*, vol. 32, no. 3. EuroVis, 2013.
- [79] W. Crozier, "The psychology of colour preferences," *Review of Progress in Coloration and Related Topics*, vol. 26, no. 1, pp. 63–72, 1996.
- [80] L.-C. Ou, M. R. Luo, A. Woodcock, and A. Wright, "A study of colour emotion and colour preference. part III: Colour preference modeling," *Color Research & Application*, vol. 29, no. 5, pp. 381–389, 2004.
- [81] J. Hutchings, "Colour in folklore and traditionthe principles," *Color Research & Application*, vol. 29, no. 1, pp. 57–66, 2004.

- [82] N. Carroll, "Art and globalization: Then and now," *The Journal of aesthetics and art criticism*, vol. 65, no. 1, pp. 131–143, 2007.
- [83] A. Ghassan and M. Blythe, "On legitimacy: designer as minor scientist," in *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2013, pp. 2149–2158.
- [84] K. Krippendorff, "On the essential contexts of artifacts or on the proposition that" design is making sense (of things)," *Design Issues*, vol. 5, no. 2, pp. 9–39, 1989.
- [85] J. Hutchings, "The continuity of colour, design, art, and science. I. the philosophy of the total appearance concept and image measurement," *Color Research & Application*, vol. 20, no. 5, pp. 296–306, 1995.
- [86] J. Hogg, S. Goodman, T. Porter, B. Mikellides, and D. Preddy, "Dimensions and determinants of judgements of colour samples and a simulated interior space by architects and non-architects," *British Journal of Psychology*, vol. 70, no. 2, pp. 231–242, 1979.
- [87] J. L. Caivano, "Research on color in architecture and environmental design: Brief history, current developments, and possible future," *Color Research & Application*, vol. 31, no. 4, pp. 350–363, 2006.
- [88] S. E. Ural and S. Yilmazer, "The architectural colour design process: An evaluation of sequential media via semantic ratings," *Color Research & Application*, vol. 35, no. 5, pp. 343–351, 2010.
- [89] T. Poldma, "Learning the dynamic processes of color and light in interior design," *Journal of Interior Design*, vol. 34, no. 2, pp. 19–33, 2009.
- [90] Y.-J. Lee and J. Lee, "The development of an emotion model based on colour combinations," *International Journal of Consumer Studies*, vol. 30, no. 2, pp. 122–136, 2006.
- [91] T. J. Madden, K. Hewett, and M. S. Roth, "Managing images in different cultures: a cross-national study of color meanings and preferences," *Journal of international marketing*, pp. 90–107, 2000.
- [92] M. Hassenzahl, "The interplay of beauty, goodness, and usability in interactive products," *Human-Computer Interaction*, vol. 19, no. 4, pp. 319–349, 2004.
- [93] A. De Angeli, A. Sutcliffe, and J. Hartmann, "Interaction, usability and aesthetics: what influences users' preferences?" in *Proceedings of the 6th conference on Designing Interactive systems*. ACM, 2006, pp. 271–280.
- [94] K. E. Schmidt, Y. Liu, and S. Sridharan, "Webpage aesthetics, performance and usability: Design variables and their effects," *Ergonomics*, vol. 52, no. 6, pp. 631–643, 2009.
- [95] D. Navarre, P. Palanque, J.-F. Ladry, and E. Barboni, "Icos: A model-based user interface description technique dedicated to interactive systems addressing usability, reliability and scalability," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 16, no. 4, p. 18, 2009.

- [96] S. Lee and R. J. Koubek, "The effects of usability and web design attributes on user preference for e-commerce web sites," *Computers in Industry*, vol. 61, no. 4, pp. 329–341, 2010.
- [97] K. Reinecke and A. Bernstein, "Improving performance, perceived usability, and aesthetics with culturally adaptive user interfaces," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 18, no. 2, p. 8, 2011.
- [98] J. Sauer and A. Sonderegger, "The influence of product aesthetics and user state in usability testing," *Behaviour & Information Technology*, vol. 30, no. 6, pp. 787–796, 2011.
- [99] K. Reinecke, T. Yeh, L. Miratrix, R. Mardiko, Y. Zhao, J. Liu, and K. Z. Gajos, "Predicting users' first impressions of website aesthetics with a quantification of perceived visual complexity and colorfulness," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2013, pp. 2049–2058.
- [100] G. Lindgaard, "Aesthetics, visual appeal, usability and user satisfaction: What do the user's eyes tell the user's brain?" *Australian Journal of Emerging Technologies & Society*, vol. 5, no. 1, 2007.
- [101] M. Hassenzahl and A. Monk, "The inference of perceived usability from beauty," *Human-Computer Interaction*, vol. 25, no. 3, pp. 235–260, 2010.
- [102] S. Lee and R. J. Koubek, "The impact of cognitive style on user preference based on usability and aesthetics for computer-based systems," *International Journal of Human-Computer Interaction*, vol. 27, no. 11, pp. 1083–1114, 2011.
- [103] T. Yang, J. Linder, and D. Bolchini, "Deep: design-oriented evaluation of perceived usability," *International Journal of Human-Computer Interaction*, vol. 28, no. 5, pp. 308–346, 2012.
- [104] M. Moshagen, J. Musch, and A. S. Göritz, "A blessing, not a curse: Experimental evidence for beneficial effects of visual aesthetics on performance," *Ergonomics*, vol. 52, no. 10, pp. 1311–1320, 2009.
- [105] A. Sonderegger and J. Sauer, "The influence of design aesthetics in usability testing: Effects on user performance and perceived usability," *Applied ergonomics*, vol. 41, no. 3, pp. 403–410, 2010.
- [106] S. Pengnate and P. Antonenko, "A multimethod evaluation of online trust and its interaction with metacognitive awareness: An emotional design perspective," *International Journal of Human-Computer Interaction*, vol. 29, no. 9, pp. 582–593, 2013.
- [107] S. Lok and S. Feiner, "A survey of automated layout techniques for information presentations," in *Proceedings of Smart Graphics*, vol. 2001, 2001, pp. 61–68.
- [108] K. Gajos and D. S. Weld, "Supple: automatically generating user interfaces," in *Proceedings of the 9th international conference on Intelligent user interfaces*. ACM, 2004, pp. 93–100.
- [109] N. Hurst, W. Li, and K. Marriott, "Review of automatic document formatting," in *Proceedings of the 9th ACM Symposium on Document Engineering (DocEng'09)*, 2009, pp. 99–108.

- [110] A. Jahanian, J. Liu, D. Tretter, Q. Lin, N. Damera-Venkata, E. O'Brien-Strain, S. Lee, J. Fan, and J. Allebach, "Automatic design of magazine covers," in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2012, pp. 83 020N–83 020N–8.
- [111] A. Hunter, D. Slatter, and D. Greig, "Web-based magazine design for self publishers," in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2011, pp. 787 902–787 902–6.
- [112] M. Kuhna, I.-M. Kivelä, and P. Oittinen, "Semi-automated magazine layout using content-based image features," in *Proceedings of the 20th ACM international conference on Multimedia*, ser. MM '12. ACM, 2012, pp. 379–388.
- [113] B. Hartmann, L. Yu, A. Allison, Y. Yang, and S. R. Klemmer, "Design as exploration: creating interface alternatives through parallel authoring and runtime tuning," in *Proceedings of the 21st annual ACM symposium on User interface software and technology*. ACM, 2008, pp. 91–100.
- [114] S. R. Herring, C.-C. Chang, J. Krantzler, and B. P. Bailey, "Getting inspired!: understanding how and why examples are used in creative design practice," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2009, pp. 87–96.
- [115] B. Lee, S. Srivastava, R. Kumar, R. Brafman, and S. R. Klemmer, "Designing with interactive example galleries," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2010, pp. 2257–2266.
- [116] R. Kumar, J. O. Talton, S. Ahmad, and S. R. Klemmer, "Bricolage: example-based retargeting for web design," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2011, pp. 2197–2206.
- [117] J. Talton, L. Yang, R. Kumar, M. Lim, N. Goodman, and R. Měch, "Learning design patterns with bayesian grammar induction," in *Proceedings of the 25th annual ACM symposium on User interface software and technology*. ACM, 2012, pp. 63–74.
- [118] J. M. Carroll, "Human–computer interaction: Psychology as a science of design," *International Journal of Human-Computer Studies*, vol. 46, no. 4, pp. 501–522, 1997.
- [119] D. Fallman, "Design-oriented human-computer interaction," in *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2003, pp. 225–232.
- [120] J. Löwgren, "Pliability as an experiential quality: Exploring the aesthetics of interaction design," *Artifact*, vol. 1, no. 2, pp. 85–95, 2007.
- [121] J. Zimmerman, J. Forlizzi, and S. Evenson, "Research through design as a method for interaction design research in hci," in *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2007, pp. 493–502.
- [122] J. Forlizzi, J. Zimmerman, and S. Evenson, "Crafting a place for interaction design research in hci," *Design Issues*, vol. 24, no. 3, pp. 19–29, 2008.
- [123] E. Stolterman, "The nature of design practice and implications for interaction design research," *International Journal of Design*, vol. 2, no. 1, pp. 55–65, 2008.

- [124] W. N. W. Hashim, N. L. M. Noor, and W. A. W. Adnan, "The design of aesthetic interaction: towards a graceful interaction framework," in *Proceedings of the 2nd International Conference on Interaction Sciences: Information Technology, Culture and Human*. ACM, 2009, pp. 69–75.
- [125] Y. Chen, "Interface and interaction design for group and social recommender systems," in *Proceedings of the fifth ACM conference on Recommender systems*. ACM, 2011, pp. 363–366.
- [126] M. Hassenzahl and N. Tractinsky, "User experience-a research agenda," *Behaviour & Information Technology*, vol. 25, no. 2, pp. 91–97, 2006.
- [127] S. Mahlke, "Visual aesthetics and the user experience," *Proceedings: The Study of Visual Aesthetics in Human-Computer Interaction*, no. 08292, 2008.
- [128] J. Hartmann, A. De Angeli, and A. Sutcliffe, "Framing the user experience: information biases on website quality judgement," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2008, pp. 855–864.
- [129] J. Löwgren, "Toward an articulation of interaction esthetics," *New Review of Hypermedia and Multimedia*, vol. 15, no. 2, pp. 129–146, 2009.
- [130] E. L.-C. Law, "The measurability and predictability of user experience," in *Proceedings of the 3rd ACM SIGCHI symposium on Engineering interactive computing systems*. ACM, 2011, pp. 1–10.
- [131] J. A. Bargas-Avila and K. Hornbæk, "Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2011, pp. 2689–2698.
- [132] P. van Schaik, M. Hassenzahl, and J. Ling, "User-experience from an inference perspective," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 19, no. 2, p. 11, 2012.
- [133] M. R. Luo, "Applying colour science in colour design," *Optics & laser technology*, vol. 38, no. 4, pp. 392–398, 2006.
- [134] H.-C. Tsai and J.-R. Chou, "Automatic design support and image evaluation of two-coloured products using colour association and colour harmony scales and genetic algorithm," *Computer-Aided Design*, vol. 39, no. 9, pp. 818–828, 2007.
- [135] L. Wang, J. Giesen, K. T. McDonnell, P. Zolliker, and K. Mueller, "Color design for illustrative visualization," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 14, no. 6, pp. 1739–1754, 2008.
- [136] L.-C. Ou, M. R. Luo, and G. Cui, "A colour design tool based on empirical studies," 2009.
- [137] S.-W. Hsiao, C.-F. Hsu, and K.-W. Tang, "A consultation and simulation system for product color planning based on interactive genetic algorithms," *Color Research & Application*, vol. 38, no. 5, pp. 375–390, 2013.

- [138] A. Jahanian, J. Liu, Q. Lin, D. R. Tretter, E. O'Brien-Strain, S. Lee, N. Lyons, and J. P. Allebach, "Automatic design of colors for magazine covers," in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2013, pp. 86 640B–86 640B.
- [139] A. Dearden and J. Finlay, "Pattern languages in hci: A critical review," *Human-computer interaction*, vol. 21, no. 1, pp. 49–102, 2006.
- [140] Y. Liu, D. Zhang, G. Lu, and W.-Y. Ma, "A survey of content-based image retrieval with high-level semantics," *Pattern Recognition*, vol. 40, no. 1, pp. 262–282, 2007.
- [141] J. Eakins, M. Graham, and T. Franklin, "Content-based image retrieval," in *Library and Information Briefings*. Citeseer, 1999.
- [142] I. K. Sethi, I. L. Coman, and D. Stan, "Mining association rules between low-level image features and high-level concepts," *Proceedings of the SPIE Data Mining and Knowledge Discovery*, vol. 3, pp. 279–290, 2001.
- [143] A. Mojsilovic and B. Rogowitz, "Capturing image semantics with low-level descriptors," in *Image Processing, 2001. Proceedings. 2001 International Conference on*, vol. 1. IEEE, 2001, pp. 18–21.
- [144] X. S. Zhou and T. S. Huang, "Cbir: from low-level features to high-level semantics," in *Electronic Imaging*. International Society for Optics and Photonics, 2000, pp. 426–431.
- [145] A. W. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 12, pp. 1349–1380, 2000.
- [146] Y. Chen, J. Z. Wang, and R. Krovetz, "An unsupervised learning approach to content-based image retrieval," in *Signal Processing and Its Applications, 2003. Proceedings. Seventh International Symposium on*, vol. 1. IEEE, 2003, pp. 197–200.
- [147] G. Csurka, S. Skaff, L. Marchesotti, and C. Saunders, "Learning moods and emotions from color combinations," in *Proceedings of the Seventh Indian Conference on Computer Vision, Graphics and Image Processing*. ACM, 2010, pp. 298–305.
- [148] "COLOURlovers," <http://www.colourlovers.com/>.
- [149] P. O'Donovan, A. Agarwala, and A. Hertzmann, "Color compatibility from large datasets," *ACM Transactions on Graphics (TOG)*, vol. 30, no. 4, p. 63, 2011.
- [150] M. Solli and R. Lenz, "Color semantics for image indexing," in *CGIV 2010: 5th European Conference on Colour in Graphics, Imaging, and Vision*. IS&T, 2010, pp. 353–358.
- [151] A. Jahanian, J. Liu, Q. Lin, D. Tretter, E. O'Brien-Strain, S. C. Lee, N. Lyons, and J. Allebach, "Recommendation system for automatic design of magazine covers," in *Proceedings of the 2013 international conference on Intelligent user interfaces*. ACM, 2013, pp. 95–106.
- [152] N. Murray, S. Skaff, L. Marchesotti, and F. Perronnin, "Toward automatic and flexible concept transfer," *Computers & Graphics*, vol. 36, no. 6, pp. 622–634, 2012.

- [153] R. Kumar, A. Satyanarayan, C. Torres, M. Lim, S. Ahmad, S. R. Klemmer, and J. O. Talton, “Webzeitgeist: design mining the web,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2013, pp. 3083–3092.
- [154] A. Mojsilovic, “A computational model for color naming and describing color composition of images,” *Image Processing, IEEE Transactions on*, vol. 14, no. 5, pp. 690–699, 2005.
- [155] J. Van De Weijer, C. Schmid, J. Verbeek, and D. Larlus, “Learning color names for real-world applications,” *Image Processing, IEEE Transactions on*, vol. 18, no. 7, pp. 1512–1523, 2009.
- [156] B. Schauerte and R. Stiefelhagen, “Learning robust color name models from web images,” in *Pattern Recognition (ICPR), 2012 21st International Conference on*. IEEE, 2012, pp. 3598–3601.
- [157] D. M. Blei, “Probabilistic topic models,” *Communications of the ACM*, vol. 55, no. 4, pp. 77–84, Apr. 2012.
- [158] Y. Feng and M. Lapata, “Topic models for image annotation and text illustration,” in *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*. Association for Computational Linguistics, 2010, pp. 831–839.
- [159] J. Heer and M. Bostock, “Crowdsourcing graphical perception: using mechanical turk to assess visualization design,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2010, pp. 203–212.
- [160] K. Reinecke and K. Z. Gajos, “Quantifying visual preferences around the world,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2014.
- [161] D. Govindaraj, T. Wang, and S. Vishwanathan, “Modeling attractiveness and multiple clicks in sponsored search results,” *arXiv preprint arXiv:1401.0255*, 2014.
- [162] “Dewey Decimal Classification,” <http://www.oclc.org/dewey/>.
- [163] “WorldCat,” <http://www.worldcat.org/>.
- [164] J. Harel, C. Koch, and P. Perona, “Graph-based visual saliency,” in *Advances in Neural Information Processing Systems 19*. MIT Press, 2007, pp. 545–552.
- [165] P. F. Felzenszwalb and D. P. Huttenlocher, “Efficient graph-based image segmentation,” *International Journal of Computer Vision*, vol. 59, no. 2, pp. 167–181, 2004.
- [166] M. F. Porter, “An algorithm for suffix stripping,” *Program: Electronic Library and Information Systems*, vol. 14, no. 3, pp. 130–137, 1980.
- [167] *Interior Design*. Sandow Media, LLC, January, 2010.
- [168] *Technology Review*. Technology Review Inc. (owned by MIT), September, 2009.
- [169] *Astronomy*. Kalmbach Publishing, January, 2010.
- [170] “Matlab Topic Modeling Toolbox,” http://psiexp.ss.uci.edu/research/programs_data/toolbox.htm.

- [171] T. L. Griffiths and M. Steyvers, “Finding scientific topics,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. Suppl 1, pp. 5228–5235, 2004.
- [172] M. Steyvers and T. Griffiths, “Probabilistic topic models,” in *Handbook of Latent Semantic Analysis*, S. D. Thomas K. Landauer, Danielle S. McNamara and W. Kintsch, Eds. Lawrence Erlbaum., 2007.
- [173] M. I. Jordan, Z. Ghahramani, T. S. Jaakkola, and L. K. Saul, “An introduction to variational methods for graphical models,” *Machine Learning*, vol. 37, no. 2, pp. 183–233, 1999.
- [174] Y. W. Teh, D. Newman, and M. Welling, “A collapsed variational Bayesian inference algorithm for latent Dirichlet allocation,” in *The Neural Information Processing Systems (NIPS)*, vol. 6, 2006, pp. 1378–1385.
- [175] C. Andrieu, N. De Freitas, A. Doucet, and M. I. Jordan, “An introduction to MCMC for machine learning,” *Machine Learning*, vol. 50, no. 1-2, pp. 5–43, 2003.
- [176] T. Griffiths, “Gibbs sampling in the generative model of latent Dirichlet allocation,” Stanford University, Tech. Rep., 2002.
- [177] I. Porteous, D. Newman, A. Ihler, A. Asuncion, P. Smyth, and M. Welling, “Fast collapsed Gibbs sampling for latent Dirichlet allocation,” in *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2008, pp. 569–577.
- [178] A. J.-B. Chaney and D. M. Blei, “Visualizing topic models.” in *The International AAAI Conference on Weblogs and Social Media (ICWSM)*, 2012.
- [179] J. Chuang, C. D. Manning, and J. Heer, “Termite: Visualization techniques for assessing textual topic models,” in *Proceedings of the International Working Conference on Advanced Visual Interfaces*. ACM, 2012, pp. 74–77.
- [180] G. Sharma, “Color fundamentals for digital imaging,” in *Digital Color Imaging Handbook*, G. Sharma, Ed. CRC Press, 2002.
- [181] H. W. Kuhn, “The Hungarian method for the assignment problem,” *Naval Research Logistics Quarterly*, vol. 2, no. 1-2, pp. 83–97, 1955.
- [182] Y. Rubner, C. Tomasi, and L. J. Guibas, “The Earth Mover’s distance as a metric for image retrieval,” *International Journal of Computer Vision*, vol. 40, no. 2, pp. 99–121, 2000.
- [183] C. Seifert, B. Kump, W. Kienreich, G. Granitzer, and M. Granitzer, “On the beauty and usability of tag clouds,” in *Information Visualisation, 2008. IV’08. 12th International Conference*. IEEE, 2008, pp. 17–25.
- [184] P. Bendito, Visual Communication Design, Patti and Rusty Rueff School of Visual and Performing Arts, Purdue University, private communication, 2013.
- [185] “Adobe Kuler,” <http://kuler.adobe.com/>.

- [186] P. O'Donovan, A. Agarwala, and A. Hertzmann, "Collaborative filtering of color aesthetics," in *Proceedings of the Workshop on Computational Aesthetics*, ser. CAe '14. New York, NY, USA: ACM, 2014, pp. 33–40. [Online]. Available: <http://doi.acm.org/10.1145/2630099.2630100>
- [187] *Interior Design*. Sandow Media, LLC, February, 2009.
- [188] B. Shneiderman, "Creativity support tools: A grand challenge for hci researchers," *Engineering the User Interface*, pp. 1–9, 2009.
- [189] S. P. Dow, A. Glassco, J. Kass, M. Schwarz, D. L. Schwartz, and S. R. Klemmer, "Parallel prototyping leads to better design results, more divergence, and increased self-efficacy," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 17, no. 4, p. 18, 2010.
- [190] C. Couprie, C. Farabet, L. Najman, and Y. LeCun, "Indoor semantic segmentation using depth information," *arXiv preprint arXiv:1301.3572*, 2013.
- [191] R. Socher, Q. Le, C. Manning, and A. Ng, "Grounded compositional semantics for finding and describing images with sentences," in *NIPS Deep Learning Workshop*, 2013.
- [192] "TripAdvisor LLC," <http://www.tripadvisor.com/>, September 20, 2014.
- [193] "The Webby Awards," <http://www.webbyawards.com/>.
- [194] S. Lin, D. Ritchie, M. Fisher, and P. Hanrahan, "Probabilistic color-by-numbers: Suggesting pattern colorizations using factor graphs," in *ACM SIGGRAPH 2013 Papers*, 2013.
- [195] B. Wang, Y. Yu, T.-T. Wong, C. Chen, and Y.-Q. Xu, "Data-driven image color theme enhancement," *ACM Transactions on Graphics (TOG)*, vol. 29, no. 6, p. 146, 2010.
- [196] A. Xu, S.-W. Huang, and B. P. Bailey, "Voyant: Generating structured feedback on visual designs using a crowd of non-experts," *Urbana*, vol. 51, p. 61801, 2014.
- [197] C. Li and T. Chen, "Aesthetic visual quality assessment of paintings," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 3, no. 2, pp. 236–252, 2009.
- [198] D. Joshi, R. Datta, E. Fedorovskaya, Q. Luong, J. Wang, J. Li, and J. Luo, "Aesthetics and emotions in images," *Signal Processing Magazine, IEEE*, vol. 28, no. 5, pp. 94–115, 2011.
- [199] M. H. Bornstein, K. Ferdinandsen, and C. G. Gross, "Perception of symmetry in infancy," *Developmental Psychology*, vol. 17, no. 1, p. 82, 1981.
- [200] V. GIANNOULI, "Visual symmetry perception," *ENCEPHALOS*, vol. 50, pp. 31–42, 2013.
- [201] S. E. Palmer, K. B. Schloss, and J. Sammartino, "Visual aesthetics and human preference," *Annual review of psychology*, vol. 64, pp. 77–107, 2013.
- [202] A. M. Norcia, T. R. Candy, M. W. Pettet, V. Y. Vildavski, and C. W. Tyler, "Temporal dynamics of the human response to symmetry," *Journal of Vision*, vol. 2, no. 2, p. 1, 2002.

- [203] A. Wilson and A. Chatterjee, "The assessment of preference for balance: Introducing a new test," *Empirical Studies of the Arts*, vol. 23, no. 2, pp. 165–180, 2005.
- [204] J. Sammartino and S. E. Palmer, "Aesthetic issues in spatial composition: Effects of vertical position and perspective on framing single objects." *Journal of Experimental Psychology: Human Perception and Performance*, vol. 38, no. 4, p. 865, 2012.
- [205] S. Gershoni and S. Hochstein, "Measuring pictorial balance perception at first glance using Japanese calligraphy," *i-Perception*, vol. 2, no. 6, p. 508, 2011.
- [206] M. H. Leyssen, S. Linsen, J. Sammartino, and S. E. Palmer, "Aesthetic preference for spatial composition in multiobject pictures," *i-Perception*, vol. 3, no. 1, p. 25, 2012.
- [207] C. M. Linnett, R. H. Morriss, W. P. Dunlap, and C. J. Fritchie, "Differences in color balance depending upon mode of comparison," *The Journal of general psychology*, vol. 118, no. 3, pp. 271–283, 1991.
- [208] U. Polat and C. W. Tyler, "What pattern the eye sees best," *Vision Research*, vol. 39, no. 5, pp. 887–895, 1999.
- [209] P. Locher, K. Overbeeke, and P. J. Stappers, "Spatial balance of color triads in the abstract art of Piet Mondrian," *Perception-London*, vol. 34, no. 2, pp. 169–190, 2005.
- [210] O. Vartanian, C. Martindale, J. Podsiadlo, S. Overbay, and J. Borkum, "The link between composition and balance in masterworks vs. paintings of lower artistic quality," *British Journal of Psychology*, vol. 96, no. 4, pp. 493–503, 2005.
- [211] P. J. Locher, P. Jan Stappers, and K. Overbeeke, "The role of balance as an organizing design principle underlying adults' compositional strategies for creating visual displays," *Acta Psychologica*, vol. 99, no. 2, pp. 141–161, 1998.
- [212] P. Locher, E. Cornelis, J. Wagemans, and P. Stappers, "Artists' use of compositional balance for creating visual displays," *Empirical Studies of the Arts*, vol. 19, no. 2, pp. 213–228, 2001.
- [213] A. W. White, *The elements of graphic design: space, unity, page architecture, and type*. Skyhorse Publishing Inc., 2002.
- [214] P. J. LOCHER, "Experimental techniques for investigating the contribution of pictorial balance to the creation and perception of visual displays," *Empirical Studies of the Arts*, vol. 21, no. 2, pp. 127–135, 2003.
- [215] K. Cullen, *Layout workbook: a real-world guide to building pages in graphic design*. Rockport Publishers, 2007.
- [216] T. Samara, *Design elements: A graphic style manual*. Rockport Publishers, 2007.
- [217] T. Samara, *Design evolution: theory into practice: a handbook of basic design principles applied in contemporary design*. Rockport Publishers, 2008.
- [218] B. Bergström, *Essentials of visual communication*. Laurence King Publishers, 2009.
- [219] E. Lupton and J. C. Phillips, *Graphic Design the New Basics*. Chronicle Books, 2014.

- [220] A. Santella, M. Agrawala, D. DeCarlo, D. Salesin, and M. Cohen, “Gaze-based interaction for semi-automatic photo cropping,” in *Proceedings of the SIGCHI conference on Human Factors in computing systems*. ACM, 2006, pp. 771–780.
- [221] L. Weitzman and K. Wittenburg, “Automatic presentation of multimedia documents using relational grammars,” in *Proceedings of the second ACM international conference on Multimedia*. ACM, 1994, pp. 443–451.
- [222] A. Borning, R. K.-H. Lin, and K. Marriott, “Constraint-based document layout for the web,” *Multimedia systems*, vol. 8, no. 3, pp. 177–189, 2000.
- [223] D. Malerba, F. Esposito, and O. Altamura, “Adaptive layout analysis of document images,” in *Foundations of Intelligent Systems*. Springer, 2002, pp. 526–534.
- [224] C. Jacobs, W. Li, E. Schrier, D. Barger, and D. Salesin, “Adaptive grid-based document layout,” *ACM transactions on graphics (TOG)*, vol. 22, no. 3, pp. 838–847, 2003.
- [225] X. Lin, “Active layout engine: Algorithms and applications in variable data printing,” *Computer-Aided Design*, vol. 38, no. 5, pp. 444–456, 2006.
- [226] J. Kustanowitz and B. Shneiderman, “Hierarchical layouts for photo libraries,” *IEEE MultiMedia*, vol. 13, no. 4, pp. 62–72, 2006.
- [227] E. Schrier, M. Dontcheva, C. Jacobs, G. Wade, and D. Salesin, “Adaptive layout for dynamically aggregated documents,” in *Proceedings of the 13th international conference on Intelligent user interfaces*. ACM, 2008, pp. 99–108.
- [228] A. Di Iorio, L. Furini, F. Vitali, J. Lumley, and T. Wiley, “Higher-level layout through topological abstraction,” in *Proceedings of the eighth ACM symposium on Document engineering*. ACM, 2008, pp. 90–99.
- [229] H. Y. Balinsky, J. R. Howes, and A. J. Wiley, “Aesthetically-driven layout engine,” in *Proceedings of the 9th ACM symposium on Document engineering*. ACM, 2009, pp. 119–122.
- [230] P. Cleveland, “Style based automated graphic layouts,” *Design Studies*, vol. 31, no. 1, pp. 3–25, 2010.
- [231] P. Merrell, E. Schkufza, Z. Li, M. Agrawala, and V. Koltun, “Interactive furniture layout using interior design guidelines,” in *ACM Transactions on Graphics (TOG)*, vol. 30, no. 4. ACM, 2011, p. 87.
- [232] M. Nebeling, F. Matulic, L. Streit, and M. C. Norrie, “Adaptive layout template for effective web content presentation in large-screen contexts,” in *Proceedings of the 11th ACM symposium on Document engineering*. ACM, 2011, pp. 219–228.
- [233] D. Raneburger, R. Popp, and J. Vanderdonckt, “An automated layout approach for model-driven wimp-ui generation,” in *Proceedings of the 4th ACM SIGCHI symposium on Engineering interactive computing systems*. ACM, 2012, pp. 91–100.
- [234] P. O’Donovan, A. Agarwala, and A. Hertzmann, “Learning layouts for single-page graphic designs,” 2014.

- [235] V. Setlur, S. Takagi, R. Raskar, M. Gleicher, and B. Gooch, “Automatic image retargeting,” in *Proceedings of the 4th international conference on Mobile and ubiquitous multimedia*. ACM, 2005, pp. 59–68.
- [236] K. Berkner and B. Erol, “Adaptation of document images to display constraints,” in *Electronic Imaging 2008*. International Society for Optics and Photonics, 2008, pp. 68 061C–68 061C.
- [237] R. Kumar, J. Kim, and S. R. Klemmer, “Automatic retargeting of web page content,” in *CHI’09 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2009, pp. 4237–4242.
- [238] L. Liu, R. Chen, L. Wolf, and D. Cohen-Or, “Optimizing photo composition,” in *Computer Graphics Forum*, vol. 29, no. 2. Wiley Online Library, 2010, pp. 469–478.
- [239] D. Vaquero, M. Turk, K. Pulli, M. Tico, and N. Gelfand, “A survey of image retargeting techniques,” in *SPIE Optical Engineering+ Applications*. International Society for Optics and Photonics, 2010, pp. 779 814–779 814.
- [240] F. Stentiford, “Attention based auto image cropping,” in *The 5th International Conference on Computer Vision Systems, Bielefeld*. Citeseer, 2007.
- [241] J. She, D. Wang, and M. Song, “Automatic image cropping using sparse coding,” in *Pattern Recognition (ACPR), 2011 First Asian Conference on*. IEEE, 2011, pp. 490–494.
- [242] S. Banerjee and B. L. Evans, “In-camera automation of photographic composition rules,” *Image Processing, IEEE Transactions on*, vol. 16, no. 7, pp. 1807–1820, 2007.
- [243] C.-Y. Lai, P.-H. Chen, S.-W. Shih, Y. Liu, and J.-S. Hong, “Computational models and experimental investigations of effects of balance and symmetry on the aesthetics of text-overlaid images,” *International journal of human-computer studies*, vol. 68, no. 1, pp. 41–56, 2010.
- [244] P. Obrador, L. Schmidt-Hackenberg, and N. Oliver, “The role of image composition in image aesthetics,” in *Image Processing (ICIP), 2010 17th IEEE International Conference on*. IEEE, 2010, pp. 3185–3188.
- [245] A. Altaboli and Y. Lin, “Investigating effects of screen layout elements on interface and screen design aesthetics,” *Advances in Human-Computer Interaction*, vol. 2011, p. 5, 2011.
- [246] P. Carpenter, “Art and ideas: An approach to art appreciation,” 1971.
- [247] F. Samuel and D. Kerzel, “Judging whether it is aesthetic: Does equilibrium compensate for the lack of symmetry?” *i-Perception*, vol. 4, no. 1, p. 57, 2013.
- [248] “500px,” <http://500px.com/>.
- [249] R. Padovan, *Proportion: science, philosophy, architecture*. Taylor & Francis, 1999.
- [250] G. T. Fechner, “Various attempts to establish a basic form of beauty: Experimental aesthetics, golden section, and square,” *Empirical studies of the arts*, vol. 15, no. 2, pp. 115–130, 1997.

- [251] H. Höge, “The golden section hypothesis—its last funeral,” *Empirical Studies of the Arts*, vol. 15, no. 2, pp. 233–255, 1997.
- [252] V. J. Konecni, “The golden section: Elusive, but detectable,” *Creativity Research Journal*, vol. 15, no. 2-3, pp. 267–275, 2003.
- [253] I. McManus, R. Cook, and A. Hunt, “Beyond the golden section and normative aesthetics: why do individuals differ so much in their aesthetic preferences for rectangles?” *Psychology of Aesthetics, Creativity, and the Arts*, vol. 4, no. 2, p. 113, 2010.
- [254] S. A. Amirshahi, G. U. Hayn-Leichsenring, J. Denzler, and C. Redies, “Evaluating the rule of thirds in photographs and paintings.”
- [255] P. J. Locher and P. J. Stappers, “Factors contributing to the implicit dynamic quality of static abstract designs,” *PERCEPTION-LONDON-*, vol. 31, no. 9, pp. 1093–1108, 2002.
- [256] C. M. Bishop *et al.*, *Pattern recognition and machine learning*. Springer New York, 2006, vol. 1.
- [257] R. Arnheim, *The power of the center: A study of composition in the visual arts*. Univ. of California Press, 1983.
- [258] J. Allebach, M. Segovia, C. Atkins, E. O’Brien-Strain, N. Damera-Venkata, N. Bhatti, J. Liu, and Q. Lin, “An investigation of document aesthetics for web-to-print repurposing of small-medium business marketing collateral,” in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 7540, 2010, p. 13.
- [259] K.-W. Koh and E. Chatow, “Magcloud: magazine self-publishing for the long tail,” in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2010, pp. 75 400E–75 400E.
- [260] A. Jahanian, “Automatic magazine cover design,” Master’s thesis, Purdue University, 2011.
- [261] L. Koren and R. W. Meckler, *Graphic design cookbook: mix & match recipes for faster, better layouts*. Chronicle books, 2001.
- [262] J. Leslie, *magCulture: new magazine design*. Laurence King, 2003.
- [263] T. Samara, *Typography: A Real-World Guide to Using Type in Graphic Design*. Rockport Pub, 2004.
- [264] T. Samara, *Making and breaking the grid: A graphic design layout workshop*. Rockport Publishers, 2005.
- [265] R. Williams, *The non-designer’s design book*. Peachpit Press, 2008.
- [266] F. D. Ching, *Architecture: Form, space, and order*. Wiley, 2010.
- [267] T. Samara, *Graphic Designer’s Essential Reference: Visual Elements, Techniques, and Layout Strategies for Busy Designers*. Rockport Publishers, 2011.

- [268] J. Felici, *The Complete Manual of typography: A Guide to setting perfect type*. Adobe Press, 2011.
- [269] D. Puhalla, *Design Elements, Form & Space: A Graphic Style Manual for Understanding Structure and Design*. Rockport Publishers, 2011.
- [270] J. Craig, W. Bevington, and I. K. Scala, *Designing with type: the essential guide to typography*. Random House LLC, 2012.
- [271] R. Carter, B. Day, and P. B. Meggs, *Typographic design: Form and communication*. Wiley, 2012.
- [272] K. Golombisky and R. Hagen, *White space is not your enemy: A beginner's guide to communicating visually through graphic, web & multimedia design*. Taylor & Francis, 2013.
- [273] *Indianapolis Monthly*. Emmis Communications, February, 2011.
- [274] X. Zhang, Y. Gao, C. B. Atkins, P. Cheatle, J. Xiao, H. Chao, P. Wu, D. Tretter, D. Slatter, A. Carter, R. Penny, and C. Willis, "Autophotobook: using technology to streamline photobook creation," in *Proceedings of the SPIE*, ser. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, vol. 7540, 2010, pp. 75 400L–75 400L–15.
- [275] D. Knuth and M. Plass, "Breaking paragraphs into lines," *Software: Practice and Experience*, vol. 11, no. 11, pp. 1119–1184, 1981.
- [276] X. Lin, "Intelligent content fitting for digital publishing," in *Electronic Imaging 2006*. International Society for Optics and Photonics, 2006, pp. 60 760J–60 760J.
- [277] T. Samara, *Typography workbook: A real-world guide to using type in graphic design*. Rockport Publishers, 2006.
- [278] J. Geigel and A. C. Loui, "Automatic page layout using genetic algorithms for electronic albuming," in *Photonics West 2001-Electronic Imaging*. International Society for Optics and Photonics, 2000, pp. 79–90.
- [279] M. V. O. Segovia, N. Damera-Venkata, E. O'Brien-Strain, J. Fan, S. H. Lim, S. Liu, J. Liu, Q. Lin, and J. P. Allebach, "Aesthetic role of transparency and layering in the creation of photo layouts," in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2011, pp. 78 790B–78 790B.
- [280] R. Poynor, *No more rules: graphic design and postmodernism*. Yale University Press, 2003.
- [281] J. Stockton, *Designer's Guide to Color 1*. Chronicle Books, 1984, vol. 1.
- [282] J. Allen, *Designer's guide to color 3*. Chronicle Books, 1986, no. 3.
- [283] I. Shibukawa and Y. Takahashi, *Designer's Guide to Color 4*. Chronicle Books, 2006, vol. 4.
- [284] H. Chijiwa, *Color harmony: a guide to creative color combinations*. Rockport Publishers, 1987.

- [285] S. Zuffi, C. Brambilla, G. Beretta, and P. Scala, “Human computer interaction: Legibility and contrast,” in *Image Analysis and Processing, 2007. ICIAP 2007. 14th International Conference on*. IEEE, 2007, pp. 241–246.
- [286] “HP MagCloud,” <http://www.magcloud.com/>.
- [287] D. Greig, A. Hunter, and D. Slatter, “Layout hierarchies for interactive design reuse,” in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 8302, 2012, p. 12.
- [288] G. Adomavicius and A. Tuzhilin, “Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions,” *Knowledge and Data Engineering, IEEE Transactions on*, vol. 17, no. 6, pp. 734–749, 2005.
- [289] “amanaimages EVE,” <http://amanaimages.com/eve/>.
- [290] E. Sparling and S. Sen, “Rating: how difficult is it?” in *Proceedings of the Fifth ACM Conference on Recommender Systems*. ACM, 2011, pp. 149–156.
- [291] R. Tibshirani, “Regression shrinkage and selection via the lasso,” *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288, 1996.
- [292] D. Frohlich, “Beauty as a design prize,” *Human–Computer Interaction*, vol. 19, no. 4, pp. 359–366, 2004.
- [293] R. Barthes, *Camera lucida: Reflections on photography*. Hill and Wang, 1982.
- [294] K. Swearingen and R. Sinha, “Interaction design for recommender systems,” in *Designing Interactive Systems*, vol. 6, no. 12. ACM Press, 2002, pp. 312–334.
- [295] B. Shneiderman, *Leonardo’s laptop: human needs and the new computing technologies*. MIT Press, 2003.
- [296] B. Carpenter, “Integrating out multinomial parameters in latent Dirichlet allocation and naïve Bayes for collapsed Gibbs sampling,” Technical report, LingPipe, Tech. Rep., 2010.

APPENDICES

A. APPENDIX DESIGN MINING COLOR SEMANTICS

Table A.1.: Summary of our magazine covers dataset. In total 2,645 magazine covers from 71 different titles were collected. These titles are roughly assigned to 12 different genres of magazines by us for the purpose of our analysis.

Art		Business	
Magazine Title	# collected	Magazine Title	# collected
Art in America	49	Bloomberg Businessweek	50
ARTNews	50	Entrepreneur	52
HOW	41	Forbes	50
ID	1	Harvard Business Review	50
Interior Design	50	Money	48
Live Design	9	The Economist	29
The New Yorker	50		
Art Total:	250	Business Total:	279
Education		Entertainment	
Magazine Title	# collected	Magazine Title	# collected
Academe	40	Conde Nast Traveler	50
Language	36	Jet	50
Language Arts	2	National Geographic	44
Social Studies and the Young Learner	50	People	50
Teaching Exceptional Children	2	Rolling Stone	50
Techniques	48		
The American Scholar	3		

continued on next page

Table A.1.: *continued*

The Reading Teacher	3		
The Science Teacher	50		
Education Total:	234	Entertainment Total:	244
Family		Fashion	
Magazine Title	# collected	Magazine Title	# collected
Good Housekeeping	61	Esquire	11
Better Homes and Gardens	30	Essence	50
House Beautiful	50	Glamour	50
Parenting	51	GQ	50
ShopSmart	40	Vogue	50
Family Total:	232	Fashion Total:	211
Health		Nature	
Magazine Title	# collected	Magazine Title	# collected
Health	17	BBC Wildlife	25
Men's Health	50	Garden Design	35
Natural Health	17	House & Garden	15
Women's Health	50	Organic Gardening	17
Health Total:	134	Nature Total:	92
Politics		Science	
Magazine Title	# collected	Magazine Title	# collected
Human Rights	42	astronomy	10
Newsweek	50	Nature	50
The Atlantic	25	Popular Science	29
The Progressive	22	Science	119
Time	50	Science and Children	5

continued on next page

Table A.1.: *continued*

		Science Illustrated	30
		Smithsonian	38
Politics Total:	189	Science Total:	281
Sports		Technology	
Magazine Title	# collected	Magazine Title	# collected
Bicycling	22	Aviation Week	45
Car and Driver	29	PC World	50
Golf Digest	26	Popular Mechanics	34
Golf World	24	Technology Review	17
Runner's World	22	Wired	45
Sporting News	44		
Sports Illustrated	50		
Sports Total:	217	Technology Total:	191

Table A.2.: Stop word list. The list of words that we exclude from the words on the covers to. Note that these words are inspected by manually visiting the first 30 words in the word topics inferred by the model.

day	inside	meaning	reveals	things	winning	minute
dos	issue	meet	romney	tips	frances	colors
double	jennifer	minutes	ryan	today	small	grows
easy	joe	month	share	top	japan	green
essence	johns	nation	shows	trick	china	simple
exclusive	julie	needed	special	ultimate	stuff	autumn
eye	kate	obama	steve	undos	pages	design
faces	klein	ons	stop	ups	mitt	
falling	losing	picks	stories	wanted	week	
free	lost	preview	summer	ways	i	
good	makeover	rated	takes	white	super	
great	making	real	talking	work	perfect	
guide	matter	red	tells	year	spring	



Fig. A.1.: Relevance matrices. (a) and (b) females, (c) and (d) males, (e) and (f) Non-US participants, (g) and (h) US participants, (i) and (j) Non-designers, (k) and (l) designers.

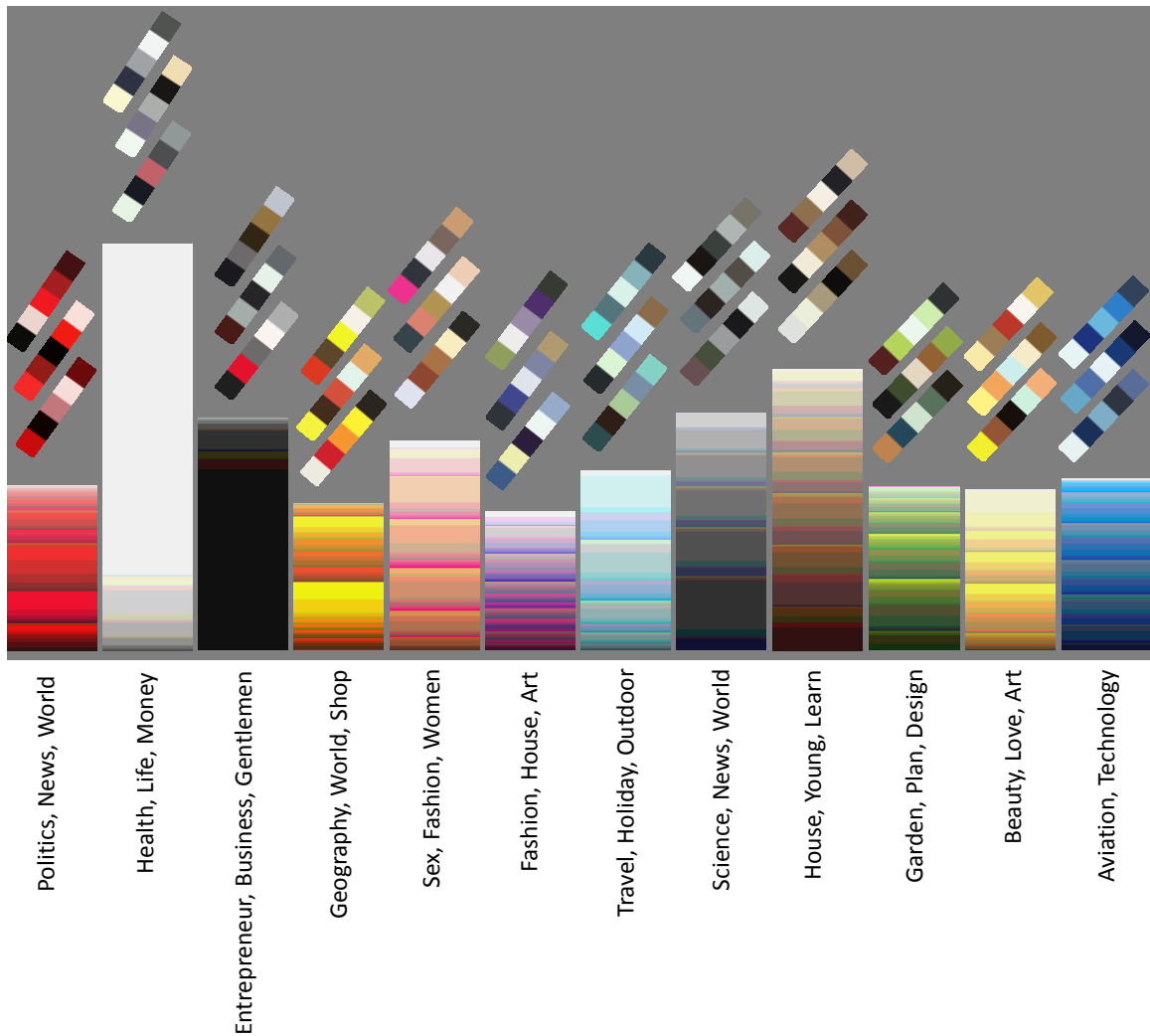
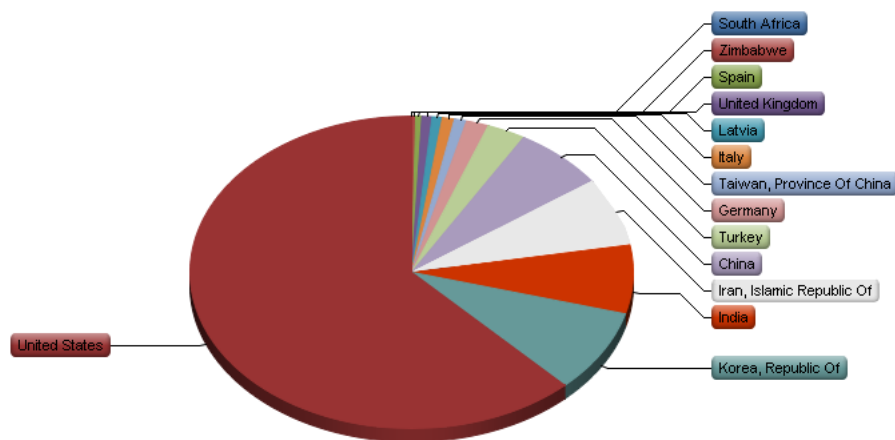
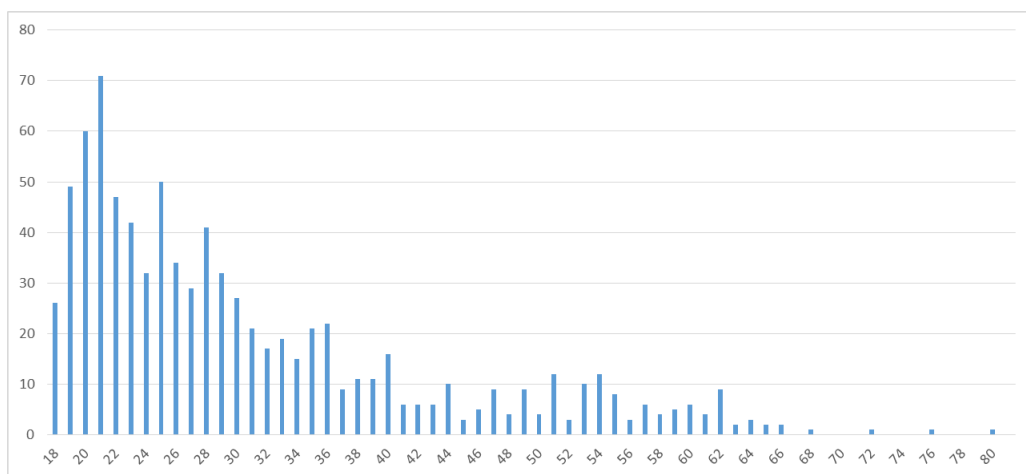


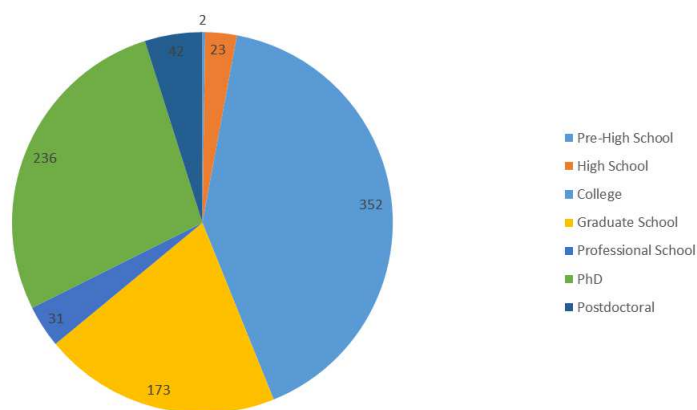
Fig. A.2.: Formative study. For this study one 5-color palette and the 12 word triples (under the color histograms) along with the “None of the above” option were shown. The participants were asked to choose as many as word triples that apply for the given 5-color palette. For each color histogram the first three closest 5-color palettes are shown in order, from top to bottom. Also for each color histogram a word triple as the representative of each word topic is shown.



(a)



(b)



(c)

Fig. A.3.: Demography statistics of the summative user study. a) 14 highest participants' countries. b) Participants' age. c) Participants' education.

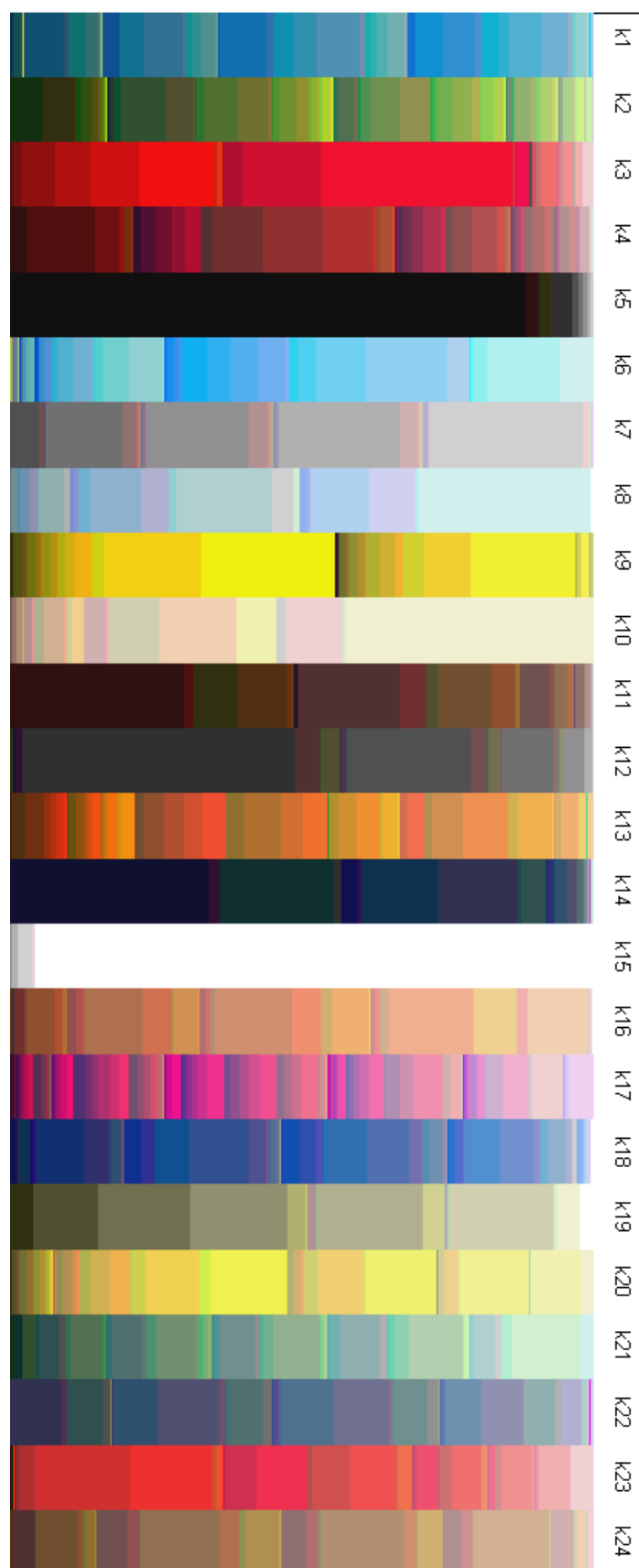


Fig. A.4.: Color topics histograms for $K = 24$.

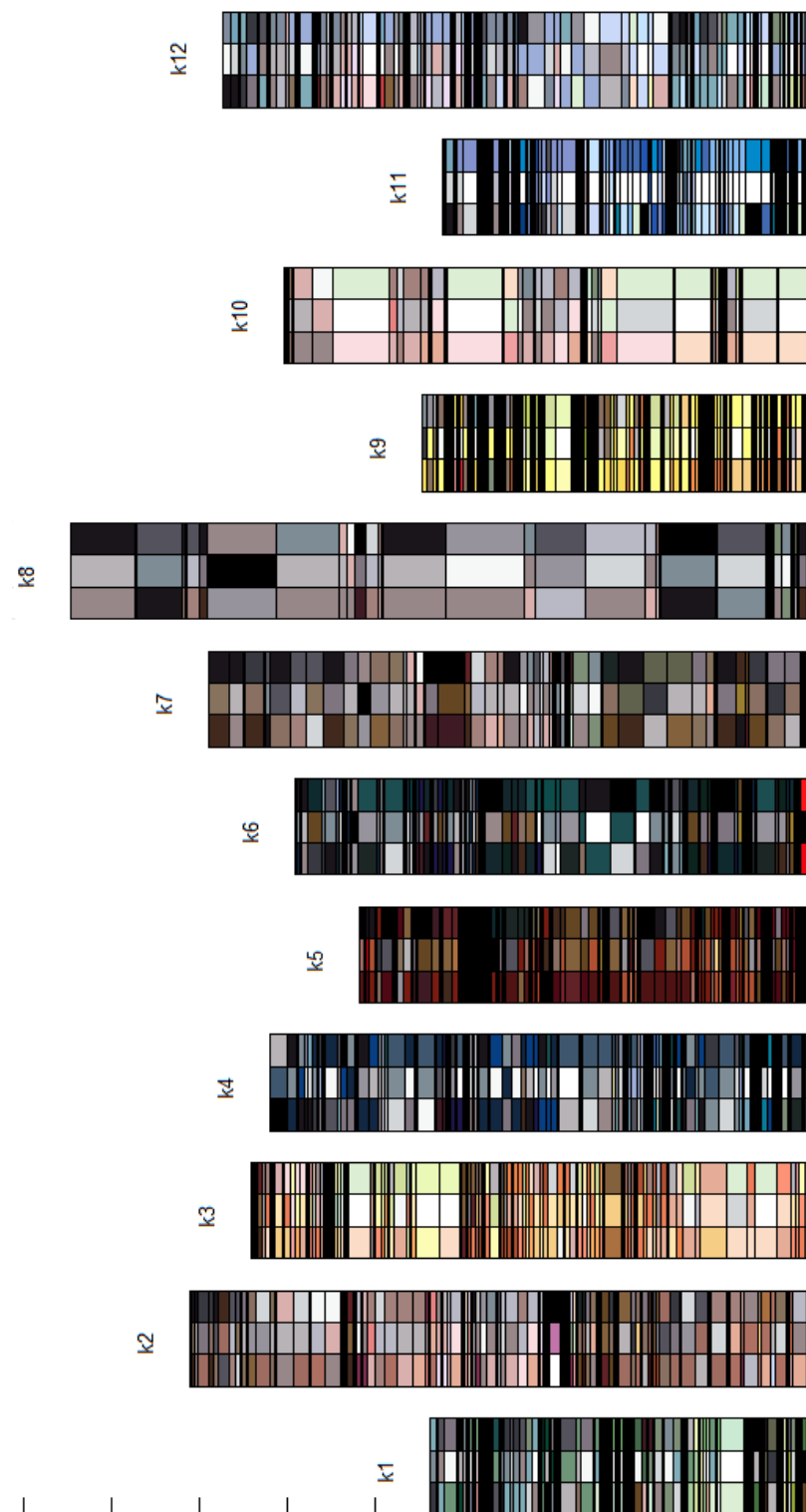


Fig. A.5.: Color topics histograms when the color basis is the 3-color combinations designed by Kobayashi [43]. Note that the black bandings are the result of small proportions. That is, when one zooms in on the visualization software tool, he/she can see these small proportions of the 3-color combinations.

B. APPENDIX DERIVATIONS OF GIBBS SAMPLING FOR LDA-DUAL

The following derivations to extend LDA is independently done by us. However, since later we found that the exact model has been earlier proposed by Shu et al. [56] called LDA-dual, although for a complete different problem, throughout the entire of this thesis, we give the credit to them and mention that we “adapt” their model. Our notation and derivation closely follows the steps in [296].

B.1 Notation

- Number of documents (magazine covers) is D .
- Number of color-word topics is K .
- Number of words in document d is $N_d \in \mathbb{N}_+$.
- Number of color swatches in document d is $M_d \in \mathbb{N}_+$.
- Number of words in the vocabulary is W .
- Number of colors in the color basis is C . For sRGB $C = 512$.
- n -th word in document d is $w_{d,n} \in \{1, \dots, W\}$.
- m -th color in document d is $c_{d,m} \in \{1, \dots, C\}$.
- The word topic to which word n in document d is assigned is $z_{d,n} \in \{1, \dots, K\}$.
- The color topic to which color m in document d is assigned is $y_{d,m} \in \{1, \dots, K\}$.
- The topic distribution for document d is $\theta_d \in \Delta_K$, where Δ_K denotes the K dimensional simplex.
- The word distribution for word topic k is $\phi_k \in \Delta_W$.
- The color distribution for color topic k is $\psi_k \in \Delta_C$.

- We will assume that prior on the word topics is the Dirichlet distribution parameterized by a single scalar β .
- Similarly, the prior on the color topics is the Dirichlet distribution parameterized by a single scalar γ .
- The prior for θ_d is a Dirichlet distribution with parameter α .
- $B(\cdot)$ denotes the beta distribution and $\Gamma(\cdot)$ denotes the gamma distribution.

We use the following notation for the sampling. The word distribution for color-word topic k is drawn from $\phi_k \sim \text{Dirichlet}(\beta)$ for $1 \leq k \leq K$. The color-word topic distribution for document d is drawn from $\theta_d \sim \text{Dirichlet}(\alpha)$ for $1 \leq d \leq D$. The distribution of the word topic assignment $z_{d,n}$ for the n -th word in document d is drawn according to $z_{d,n} \sim \text{Discrete}(\theta_d)$ for $1 \leq d \leq D, 1 \leq n \leq N_d$. Similarly, for each color in document d , we follow the same mechanism: $y_{d,m} \sim \text{Discrete}(\theta_d)$ for $1 \leq d \leq D, 1 \leq m \leq M_d$.

B.2 Joint Probability

$$\begin{aligned}
 p(\phi, \psi, \theta, \mathbf{z}, \mathbf{c}, \mathbf{y}, \mathbf{w}, | \alpha, \beta, \gamma) = & \\
 & p(\theta | \alpha) \cdot p(\mathbf{z} | \theta) \cdot p(\mathbf{y} | \theta) \cdot p(\mathbf{w} | \mathbf{z}, \phi) \cdot \\
 & p(\mathbf{c} | \mathbf{y}, \psi) \cdot p(\phi | \beta) \cdot p(\psi | \gamma). \tag{B.1}
 \end{aligned}$$

Using standard iid assumptions on the documents in the corpus, Eq B.1 can be rewritten as

$$\begin{aligned}
 p(\phi, \psi, \theta, \mathbf{z}, \mathbf{c}, \mathbf{y}, \mathbf{w}, | \alpha, \beta, \gamma) = & \prod_{d=1}^D p(\theta_d | \alpha) \cdot \prod_{d=1}^D \prod_{n=1}^{N_d} p(z_{d,n} | \theta_d) \cdot \\
 & \prod_{d=1}^D \prod_{m=1}^{M_d} p(y_{d,m} | \theta_d) \cdot \prod_{d=1}^D \prod_{m=1}^{M_d} p(c_{d,m} | \mathbf{y}, \psi) \cdot \\
 & \prod_{d=1}^D \prod_{n=1}^{N_d} p(w_{d,n} | \mathbf{z}, \phi) \cdot \prod_{k=1}^K P(\phi_k | \beta) \cdot \prod_{k=1}^K P(\psi_k | \gamma). \tag{B.2}
 \end{aligned}$$

Using our priors, Eq B.2 is

$$\begin{aligned}
p(\boldsymbol{\phi}, \boldsymbol{\psi}, \boldsymbol{\theta}, \mathbf{z}, \mathbf{c}, \mathbf{y}, \mathbf{w}, | \alpha, \beta, \gamma) &= \prod_{d=1}^D \text{Dirichlet}(\boldsymbol{\theta}_d | \alpha) \cdot \prod_{d=1}^D \prod_{n=1}^{N_d} \text{Discrete}(z_{d,n} | \boldsymbol{\theta}_d) \cdot \\
&\prod_{d=1}^D \prod_{m=1}^{M_d} \text{Discrete}(y_{d,m} | \boldsymbol{\theta}_d) \cdot \prod_{d=1}^D \prod_{m=1}^{M_d} \text{Discrete}(c_{d,m} | \mathbf{y}, \boldsymbol{\psi}) \cdot \\
&\prod_{d=1}^D \prod_{n=1}^{N_d} \text{Discrete}(w_{d,n} | \mathbf{z}, \boldsymbol{\phi}) \cdot \prod_{k=1}^K \text{Dirichlet}(\boldsymbol{\phi}_k | \beta) \cdot \\
&\prod_{k=1}^K \text{Dirichlet}(\boldsymbol{\psi}_k | \gamma). \tag{B.3}
\end{aligned}$$

Now we introduce some counting terms: $l_{k,d,w}$ is number of times word w is assigned to color-word topic k in document d . $l_{k,*,w}$ is number of times word w is assigned to color-work topic k in all documents. $l_{k,d,*}$ is total number of words in document d assigned to color-word topic k . $l_{k,*,*}$ is total number of words in all documents assigned to color-word topic k . Similarly, for color counts, we use $r_{k,d,c}$, $r_{k,*,c}$, $r_{k,d,*}$, $r_{k,*,*}$.

$$\begin{aligned}
l_{k,d,w} &= \sum_{n=1}^{N_d} I(z_{d,n} = k, w_{d,n} = w) \\
l_{k,*,w} &= \sum_{d=1}^D l_{k,d,w} \\
l_{k,d,*} &= \sum_{w=1}^W l_{k,d,w} \\
l_{k,*,*} &= \sum_{d=1}^D \sum_{w=1}^W l_{k,d,w}.
\end{aligned}$$

B.3 Integrating Out Multinomials

In the collapsed Gibbs sampling, we walk through the samples at a time. Let $\mathbf{z}_{-(i,j)} = \mathbf{z} - z_{(i,j)}$. For sampling the goal is to compute the probability of word topic assignment

$z_{(i,j)}$ that is being assigned to $w_{(i,j)}$, the i -th word of the j -th document, given that all the other word topic assignments $\mathbf{z}_{-(i,j)}$ are assigned:

$$\begin{aligned} p(z_{i,j} \mid \mathbf{z}_{-(i,j)}, \mathbf{c}, \mathbf{y}, \mathbf{w}, \alpha, \beta, \gamma) &= \frac{p(z_{i,j}, \mathbf{z}_{-(i,j)}, \mathbf{c}, \mathbf{y}, \mathbf{w} \mid \alpha, \beta, \gamma)}{p(\mathbf{z}_{-(i,j)}, \mathbf{c}, \mathbf{y}, \mathbf{w} \mid \alpha, \beta, \gamma)} \\ &\propto p(\mathbf{z}, \mathbf{c}, \mathbf{y}, \mathbf{w} \mid \alpha, \beta, \gamma). \end{aligned} \quad (\text{B.4})$$

Now

$$\begin{aligned} p(\mathbf{z}, \mathbf{c}, \mathbf{y}, \mathbf{w} \mid \alpha, \beta, \gamma) &= \int_{\boldsymbol{\theta}} \int_{\boldsymbol{\phi}} \int_{\boldsymbol{\psi}} p(\mathbf{z}, \mathbf{c}, \boldsymbol{\theta}, \mathbf{y}, \mathbf{w}, \boldsymbol{\theta}, \boldsymbol{\phi}, \boldsymbol{\psi} \mid \alpha, \beta, \gamma) d\boldsymbol{\psi}_k d\boldsymbol{\phi}_k d\boldsymbol{\theta}_d \\ &= \int_{\boldsymbol{\theta}} \int_{\boldsymbol{\phi}} \int_{\boldsymbol{\psi}} p(\boldsymbol{\theta} \mid \alpha) \cdot p(\mathbf{z} \mid \boldsymbol{\theta}) \cdot p(\mathbf{y} \mid \boldsymbol{\theta}) \cdot p(\boldsymbol{\phi} \mid \beta) \cdot p(\mathbf{w} \mid \boldsymbol{\phi}, \mathbf{z}) \cdot \\ &\quad p(\boldsymbol{\psi} \mid \gamma) \cdot p(\mathbf{c} \mid \boldsymbol{\psi}, \gamma) d\boldsymbol{\psi}_k d\boldsymbol{\phi}_k d\boldsymbol{\theta}_d \\ &= \int_{\boldsymbol{\theta}} p(\boldsymbol{\theta} \mid \alpha) \cdot p(\mathbf{z} \mid \boldsymbol{\theta}) \cdot p(\mathbf{y} \mid \boldsymbol{\theta}) d\boldsymbol{\theta}_d \times \int_{\boldsymbol{\phi}} p(\boldsymbol{\phi} \mid \beta) \cdot p(\mathbf{w} \mid \boldsymbol{\phi}, \mathbf{z}) d\boldsymbol{\phi}_k \\ &\quad \times \int_{\boldsymbol{\psi}} p(\boldsymbol{\psi} \mid \gamma) \cdot p(\mathbf{c} \mid \boldsymbol{\psi}, \gamma) d\boldsymbol{\psi}_k. \end{aligned} \quad (\text{B.5})$$

According to Dirichlet(α) distribution

$$p(\boldsymbol{\theta} \mid \alpha) = \frac{1}{B(\alpha)} \prod_{d=1}^D \prod_{k=1}^K \theta_{d,k}^{\alpha-1}.$$

Using the discrete distribution

$$p(\mathbf{z} \mid \boldsymbol{\theta}) = \prod_{d=1}^D \prod_{k=1}^K \theta_d^{l_{k,d,*}}, \text{ and } p(\mathbf{y} \mid \boldsymbol{\theta}) = \prod_{d=1}^D \prod_{k=1}^K \theta_d^{r_{k,d,*}}.$$

Hence, the first integral in (B.5) can be written as

$$\begin{aligned} \int_{\boldsymbol{\theta}} p(\boldsymbol{\theta} \mid \alpha) \cdot p(\mathbf{z} \mid \boldsymbol{\theta}) \cdot p(\mathbf{y} \mid \boldsymbol{\theta}) d\boldsymbol{\theta}_d &= \frac{1}{B(\alpha)} \int_{\boldsymbol{\theta}} \prod_{d=1}^D \prod_{k=1}^K \theta_{d,k}^{\alpha-1} \cdot \theta_{d,k}^{l_{k,d,*}} \cdot \theta_{d,k}^{r_{k,d,*}} d\boldsymbol{\theta}_d \\ &= \frac{1}{B(\alpha)} \prod_{d=1}^D \int_{\boldsymbol{\theta}} \prod_{k=1}^K \theta_{d,k}^{\alpha+l_{k,d,*}+r_{k,d,*}-1} d\boldsymbol{\theta}_d. \end{aligned}$$

Let $\alpha + l_{k,d,*} + r_{k,d,*} = \alpha'$. Then inside the integral looks like a Dirichlet(α'), and the integral sums up to 1:

$$\begin{aligned} &= \frac{1}{B(\alpha)} \prod_{d=1}^D \int_{\theta} \frac{B(\alpha')}{B(\alpha')} \prod_{k=1}^K \theta_{d,k}^{\alpha'_k - 1} d\theta_d \\ &= \frac{1}{B(\alpha)} \prod_{d=1}^D B(\alpha')(1) \\ &= \prod_{d=1}^D \frac{B(\alpha')}{B(\alpha)}. \end{aligned}$$

In a similar fashion we solve the the second and the third integral:

$$\prod_{k=1}^K \frac{B(\beta')}{B(\beta)}, \text{ where } \beta' = l_{k,*,w} + \beta, \text{ and } \prod_{k=1}^K \frac{B(\gamma')}{B(\gamma)}, \text{ where } \gamma' = r_{k,*,c} + \gamma.$$

Hence,

$$\begin{aligned} p(\mathbf{z}, \mathbf{c}, \mathbf{y}, \mathbf{w} \mid \alpha, \beta, \gamma) &= \prod_{d=1}^D \frac{B(\alpha')}{B(\alpha)} \prod_{k=1}^K \frac{B(\beta')}{B(\beta)} \prod_{k=1}^K \frac{B(\gamma')}{B(\gamma)} \\ &= \prod_{d=1}^D \frac{\Gamma(\sum_{k=1}^K \alpha)}{\prod_{k=1}^K \Gamma(\alpha)} \frac{\prod_{k=1}^K \Gamma(\alpha + l_{k,d,*} + r_{k,d,*})}{\Gamma(\sum_{k=1}^K \alpha + l_{k,d,*} + r_{k,d,*})} \\ &\quad \prod_{k=1}^K \frac{\Gamma(\sum_{w=1}^W \beta)}{\sum_{w=1}^W \Gamma(\beta)} \frac{\prod_{w=1}^W \Gamma(l_{k,*,w} + \beta)}{\Gamma(\sum_{w=1}^W l_{k,*,w} + \beta)} \\ &\quad \prod_{k=1}^K \frac{\Gamma(\sum_{c=1}^C \gamma)}{\sum_{c=1}^C \Gamma(\gamma)} \frac{\prod_{c=1}^C \Gamma(r_{k,*,c} + \gamma)}{\Gamma(\sum_{c=1}^C \gamma + r_{k,*,c})}. \end{aligned}$$

Note that Γ functions that depend on hyperparameters are constants and we can remove them to obtain

$$\begin{aligned} p(\mathbf{z}, \mathbf{c}, \mathbf{y}, \mathbf{w} \mid \alpha, \beta, \gamma) &\propto \prod_{d=1}^D \frac{\prod_{k=1}^K \Gamma(\alpha + l_{k,d,*} + r_{k,d,*})}{\Gamma(\sum_{k=1}^K \alpha + l_{k,d,*} + r_{k,d,*})} \prod_{k=1}^K \frac{\prod_{w=1}^W \Gamma(l_{k,*,w} + \beta)}{\Gamma(\sum_{w=1}^W l_{k,*,w} + \beta)} \\ &\quad \prod_{k=1}^K \frac{\prod_{c=1}^C \Gamma(r_{k,*,c} + \gamma)}{\Gamma(\sum_{c=1}^C r_{k,*,c} + \gamma)}. \end{aligned}$$

Now, at the current position (i, j) take out the terms that depend on the sample position:

$$\begin{aligned} &= \prod_{d \neq i}^D \frac{\prod_{k=1}^K \Gamma(\alpha + l_{k,d,*} + r_{k,d,*})}{\Gamma(\sum_{k=1}^K \alpha + l_{k,d,*} + r_{k,d,*})} \frac{\prod_{k=1}^K \Gamma(\alpha + l_{k,i,*} + r_{k,i,*})}{\Gamma(\sum_{k=1}^K \alpha + l_{k,i,*} + r_{k,i,*})} \\ &\quad \prod_{k=1}^K \frac{\prod_{w \neq w_{i,j}} \Gamma(l_{k,*,w} + \beta) \Gamma(l_{k,*,w_{i,j}} + \beta)}{\Gamma(\sum_{w=1}^W l_{k,*,w} + \beta)} \prod_{k=1}^K \frac{\prod_{c=1}^C \Gamma(r_{k,*,c} + \gamma)}{\Gamma(\sum_{c=1}^C \gamma + r_{k,*,c})}. \end{aligned}$$

Next, we remove the terms that do not depend on the current sample position (i, j) :

$$\propto \frac{\prod_{k=1}^K \Gamma(\alpha + l_{k,i,*} + r_{k,i,*})}{\Gamma(\sum_{k=1}^K \alpha + l_{k,i,*} + r_{k,i,*})} \prod_{k=1}^K \frac{\Gamma(l_{k,*,w_{i,j}} + \beta)}{\Gamma(\sum_{w=1}^W l_{k,*,w} + \beta)}.$$

Denote $l^{-(i,j)}$ for the counts without position (i, j) . That is $l = l^{-(i,j)}$ and to include the next count we obtain $1 + l^{-(i,j)}$:

$$\propto \frac{\prod_{k \neq z_{i,j}}^K \Gamma(\alpha + l_{k,i,*}^{-(i,j)} + r_{k,i,*}) \Gamma(\alpha + l_{z_{i,j},i,*}^{-(i,j)} + r_{z_{i,j},i,*} + 1)}{\Gamma(1 + \sum_{k=1}^K \alpha + l_{k,i,*}^{-(i,i)})} \\ \prod_{k \neq z_{i,j}}^K \frac{\Gamma(l_{k,*,w_{i,j}}^{-(i,j)} + \beta)}{\Gamma(\sum_{w=1}^W l_{k,*,w} + \beta)} \frac{\Gamma(l_{z_{i,j}}^{-(i,j)} + \beta + 1)}{\Gamma(\sum_{w=1}^W l_{z_{i,j},*,w}^{-(i,j)} + \beta)}.$$

Using the property of Γ function, $\Gamma(x + 1) = x\Gamma(x)$, we can expand the incremented terms based on the current sample position (i, j) :

$$= \frac{\prod_{k \neq z_{i,j}}^K \Gamma(\alpha + l_{k,i,*}^{-(i,j)} + r_{k,i,*}) \Gamma(\alpha + l_{z_{i,j},i,*}^{-(i,j)} + r_{z_{i,j},i,*}) (\alpha + l_{z_{i,j},i,*}^{-(i,j)} + r_{z_{i,j},i,*})}{\Gamma(1 + \sum_{k=1}^K \alpha + l_{k,i,*}^{-(i,i)})} \\ \prod_{k \neq z_{i,j}}^K \frac{\Gamma(l_{k,*,w_{i,j}}^{-(i,j)} + \beta)}{\Gamma(\sum_{w=1}^W l_{k,*,w} + \beta)} \frac{\Gamma(l_{z_{i,j}}^{-(i,j)} + \beta) (l_{z_{i,j}}^{-(i,j)} + \beta)}{\Gamma(\sum_{w=1}^W l_{z_{i,j},*,w}^{-(i,j)} + \beta)}.$$

Now shrink Γ 's into the general form:

$$= \frac{\prod_{k=1}^K \Gamma(\alpha + l_{k,i,*}^{-(i,j)} + r_{k,i,*}) (\alpha + l_{z_{i,j},i,*}^{-(i,j)} + r_{z_{i,j},i,*})}{\Gamma(1 + \sum_{k=1}^K \alpha + l_{k,i,*}^{-(i,i)})} \prod_{k=1}^K \frac{\Gamma(l_{k,*,w_{i,j}}^{-(i,j)} + \beta)}{\Gamma(\sum_{w=1}^W l_{k,*,w} + \beta)} \frac{(l_{z_{i,j}}^{-(i,j)} + \beta)}{\sum_{w=1}^W l_{z_{i,j},*,w}^{-(i,j)} + \beta}.$$

We continue with proportionality; here Γ 's are constants so we remove them:

$$\propto \frac{(\alpha + l_{k,i,*}^{-(i,j)} + r_{k,i,*}) (l_{z_{i,j}}^{-(i,j)} + \beta)}{\sum_{w=1}^W l_{z_{i,j},*,w}^{-(i,j)} + \beta}. \quad (\text{B.6})$$

Since $\sum_{w=1}^W \beta = W\beta$, Eq B.6 can be rewritten as

$$= \frac{(\alpha + l_{k,i,*}^{-(i,j)} + r_{k,i,*}) (l_{z_{i,j}}^{-(i,j)} + \beta)}{l_{z_{i,j},*,*}^{-(i,j)} + W\beta}. \quad (\text{B.7})$$

In order to compute the probability, we normalize the so far proportional computation:

$$p(z_{i,j} \mid \mathbf{z}_{-(i,j)}, \mathbf{c}, \mathbf{y}, \mathbf{w}, \alpha, \beta, \gamma) = \frac{\left(\frac{(\alpha + l_{z_{i,j},i,*}^{-(i,j)} + r_{k,i,*}) (l_{z_{i,j},*,w_{i,j}}^{-(i,j)} + \beta)}{l_{z_{i,j},*,*}^{-(i,j)} + W\beta} \right)}{\left(\sum_{k=1}^K \frac{(\alpha + l_{k,i,*}^{-(i,j)} + r_{k,i,*}) (l_{k,*,w_{i,j}}^{-(i,j)} + \beta)}{l_{k,*,*}^{-(i,j)} + W\beta} \right)}.$$

In a similar fashion, for $c_{i,j}$ we get:

$$p(y_{i,j} \mid \mathbf{y}_{-(i,j)}, \mathbf{w}, \mathbf{z}, \mathbf{c}, \alpha, \beta, \gamma) = \frac{\left(\frac{(\alpha + r_{y_{i,j},i,*}^{-(i,j)} + l_{k,i,*})(r_{y_{i,j},*,c_{i,j}}^{-(i,j)} + \gamma)}{r_{y_{i,j},*,*}^{-(i,j)} + C\gamma} \right)}{\left(\sum_{k=1}^K \frac{(\alpha + r_{k,i,*}^{-(i,j)} + r_{k,i,*})(r_{k,*,c_{i,j}}^{-(i,j)} + \gamma)}{r_{k,*,*}^{-(i,j)} + C\gamma} \right)}.$$

θ , ϕ , and ψ are estimated from the computed \mathbf{z} and \mathbf{y} :

$$\hat{\theta}_{d,k} = \frac{l_{k,d,*} + r_{k,d,*} + \alpha}{l_{k,d,*} + r_{k,d,*} + K\alpha}, \hat{\phi}_k = \frac{l_{k,*,w} + \beta}{l_{k,*,*} + W\beta}, \hat{\psi}_k = \frac{l_{k,*,c} + \gamma}{r_{k,*,*} + C\gamma}.$$