
Assistive Technology for the *Aesthetically* Impaired

John C. Hart

Dept. of Computer Science
University of Illinois
201 N. Goodwin Ave.
Urbana, IL 61801
jch@illinois.edu

Abstract

Consumers now have the ability to express themselves through the creation of online media, but few have the training and expertise to do so effectively and aesthetically. This position paper presents a vision for the future of digital content creation software, to automatically advise its users during the media creation and editing process on how to produce more effective and compelling media. It examines the production of text, images, video and audio, and examines issues of user studies and machine learning before concluding that the codification of applied art and design expertise is likely the best choice for achieving this goal.

Introduction

Digital Content Creation (DCC) is a sector of technology and software development focusing on the creation, processing and sharing of media. It includes the digital photography, audio and video used to record media as well as the accompanying software used to process media into a form communicated to others. Market analysis has shown the DCC software market has grown from \$3B to \$4B in the past two years and will likely grow another billion over the next five years (modulo the unforeseen current economic conditions).

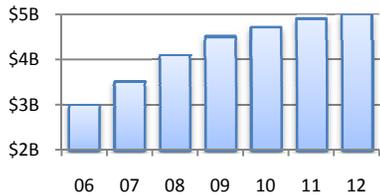


Figure 1. DCC software market growth. (Courtesy Jon Peddie Research)

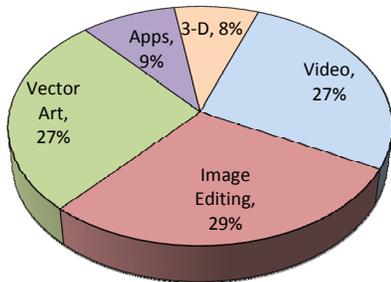


Figure 2. DCC market by content type. (Courtesy Jon Peddie Research)

DCC software development has also been fueled by the social web and its community sharing websites like YouTube, Flickr, Picasa, etc. that allow users to create and share a variety of self-produced media. These websites create a highly competitive environment for media which motivates contributors to strive to make their content as appealing as possible.

That motivation is stymied by the fact that most consumers producing and sharing media have no training formal or otherwise in media production. Consumer tools for media production and processing abound. Such software provides a powerful array of image, video and audio processing, but little if any advice on how to effectively use these tools to make one's media look better.

We consider computational creativity from the standpoints of the usability and effectiveness of its tools, and envision software that not only provides the ability to process media, but the advice needed to make the media more appealing.

DCC Areas v. Depth

We can examine features and abilities of DCC software across the media production areas of text, images, video and audio. This provides us the opportunity to document present day abilities and the abilities we can envision for future DCC software.

Text. Modern word processors provide consumers with abilities previously only available to published authors, including typesetting and proofreading for spelling and grammar errors. Recent word processors can also detect the passive voice and suggest active voice alternatives. Future word processors might provide content-dependent stylistic advice, e.g. building drama for a storytelling style, or following a journalistic style that summarizes first with details later. Such software could also examine the composition and pacing to ensure the style is employed effectively. Unfortunately such features would require natural language capabilities that far exceed current research results.

Images. Modern software provides a wealth of tools for image synthesis, processing and editing but little advice on how to use these tools to make a picture better. We envision tools that would examine a picture's composition, framing, lighting and posing and provide suggestions on alternatives to produce a picture that better communicates the author's intent. Future cameras may build in advice on composition, posing and lighting by examining the current viewfinder image. Such tools would require sophisticated vision techniques in scene understanding to decipher the content of an image before any meaningful advice could be generated.

	Text	Image	Video	Audio
HAVE:	<ul style="list-style-type: none"> Typesetting Spelling Grammar 	<ul style="list-style-type: none"> Synthesis Editing 	<ul style="list-style-type: none"> Recording Non-Linear Editing Processing 	<ul style="list-style-type: none"> Recording Playback Looping Sequencing
WANT:	<ul style="list-style-type: none"> Writing Style Composition Storytelling 	<ul style="list-style-type: none"> Cropping Composition Posing Lighting 	<ul style="list-style-type: none"> Directing Lighting Blocking Cuts/Fades 	<ul style="list-style-type: none"> Arranging Scoring Harmonic Structure

Table 1. Analysis of assistance across DCC media.

Why Expertise Codification?

"Artists command the forces of light, composition, color, and tone – whether they wield pencils and paint or lenses and cameras. As cinematographer Vittorio Storaro said, "[When] you make a design, shoot a picture, or photograph a movie, it is the representation of all 2,000 years of history, whether you are conscious of it or not.'" (Ross Lowell. *Matters of Light & Depth*. Lower Light Management, 1999.)

Band-in-a-Box's Success

Some software already supports stylistic high-level assistance for audio content. For example, given a song's chord structure and a desired style, Band-in-a-Box will provide a decent accompaniment and can even improvise a solo. This works because professional musicians were able to codify how to aesthetically connect musical phrases from a style database to work in western music forms.

Video. Modern DCC software provides tools for digital video recording and processing, as well as non-linear editing tools that enable consumers with features previously only available to professional post-production crews. Modern video editing software can automatically segment video into individual shots and provides a variety of different transitions between shots. Future DCC software might provide further analysis and advice on which transition to use between shots. Future digital video cameras might provide advice on lighting, directing and blocking to improve the content of the recorded video.

Audio. Modern consumer audio consists of tools and devices for MP3 recording and playback, plus processing tools for MP3 and MIDI looping and sequencing. "Prosumer" tools also exist for arranging, harmonizing and improvising, such as Band-in-a-Box. Future audio software might further simplify or complicate an arrangement based on a player's abilities, or might automatically generate a musical score based on an accompanying visual's content or desired mood.

Issues of Approach

The idea of analyzing content to automatically provide advice raises issues on how to approach the problem.

User Studies. A series of experiments could be conducted to discern how artists and designers go about analyzing and improving media, but art and design is well documented and such experiments would require a lot of resources to simply rediscover the same art and design principles that educated its subjects.

Machine Learning. We could discover how media is improved through the analysis of large amounts of before/after data. However, other than in isolated tutorials, the volume of before/after media examples needed for training in such an approach currently do not seem to exist. One could construct an online media improvement game [1] to provide such data. Alternatively DCC software deployed on the web could also be mined to find improvements from the input and the result. In both cases, ranking of the data would be necessary to ensure its level of expertise and authority.

Expertise Codification. The most promising approach, as demonstrated by Band-in-a-Box, appears to be the codification of expertise. This is a similar approach to other previous computational aesthetics papers, such as the virtual cinematographer which codified the rules for filming dialog [1]. This approach relies on the fact that art and design have millennia of documentation behind them, and seeks to convert its advice into an automatic form. It is worth mentioning here a preference for design and industrial art as opposed to fine art. The goal of fine art is to experiment and be cutting edge, whereas the applied arts use these results to make things look better or communicate more effectively.

Some Illustration Examples

Our current work on automatic illustration algorithms serve as an initial platform for codifying artistic expertise and providing aesthetic advice.

Mathematical Illustration. Books on calculus and other subjects in mathematics often contain beautiful illustrations of mathematical surfaces [3]. However not all mathematicians have the ability to draw a high-

Photorealism Mathematica Illustration

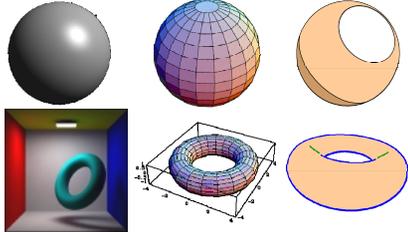


Figure 3. Illustration tools can provide better depiction of math shapes.

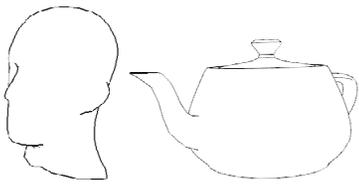


Figure 4. Three-quarter views maximize occlusions that aid perception of line drawings.

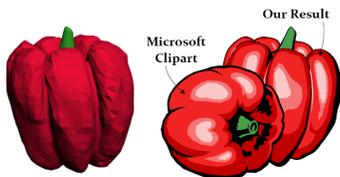


Figure 5. Mesh (left) and rendered as custom vector clip-art (right).

quality illustration. Packages like Mathematica visualize surfaces as a polygonal mesh lit by red, green and blue lights along the three coordinate axes. We have developed a tool that converts an implicit surface of the form $f(x,y,z)=0$ into a vector line drawing with filled regions that depict a shape's silhouette, shadow boundary and highlight gleams [4]. Such tools enable mathematicians to better illustrate surfaces, but do not provide advice on how to pose or light the objects.

Posing Illustrations. Much of the previous work on automatic posing has focused on maximizing the display of information, which often yields a head-on full-frontal view. Portrait artists know that a three-quarter view is better, and for line drawings serves to cue the viewer of the 3-D visibility ordering. Thus the automatic posing of illustrations seeks to maximize perceptual cues rather than information display.

Custom Clip Art. Consumers commonly make presentations and prepare documents decorated with clip-art. Clip-art is stylized drawings that use layers of 2-D closed filled polygons to depict a 3-D shape's silhouette, shadow and highlights. We have developed a rendering system that converts a 3-D mesh, such as can be downloaded from Google's 3-D warehouse or constructed using Sketchup, into a 2-D clip-art representation. We have also analyzed numerous styles of clip-art found in Microsoft Office's library, and codified these styles so they could be applied in general to 2-D line drawings [5].

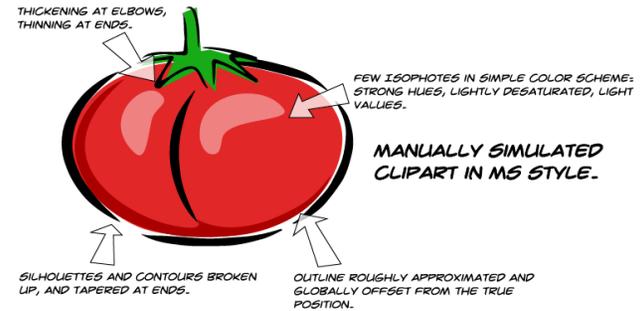


Figure 6. Codification of clip-art style by Holger Winnemoeller.

Conclusion

Hence, through the codification of applied art and design principles, we can enable consumers to use DCC software to make more effective and pleasing media.

Acknowledgments. These ideas are the result of numerous conversations and/or work with Michael Cohen, Elmar Eisemann, David Salesin, Matei Stroila, Wen Su and Holger Winnemoeller. This work has been funded by Adobe, NSF (CCR-0219594) and the INRIA/CNRS Exploradoc program.

References

- [1] L. von Ahn, R. Liu, M. Blum. Peekaboom: A Game for Locating Objects in Images. ACM Conference on Human Factors in Computing Systems, CHI 2006.
- [2] Li-wei He, Michael F. Cohen, and David Salesin. The Virtual Cinematographer: A Paradigm for Automatic Real-Time Camera Control and Directing. Proc. SIGGRAPH 96.
- [3] G. Francis. A Topological Picturebook. Springer, 1987.
- [4] M. Stroila, E. Eisemann, J.C. Hart. Clip Art Rendering of Smooth Isosurfaces. IEEE TVCG 14(1), Jan. 2008, pp. 71-81.
- [5] E. Eisemann, H. Winnemoller, J.C. Hart, D. Salesin. Stylized Vector Art from 3D Models with Region Support. Proc. Eurographics Rendering Symp. 2008.